CrossMark

Global well-posedness and scattering of the (4+1)-dimensional Maxwell-Klein-Gordon equation

Sung-Jin Oh[1] · Daniel Tataru[1]

Abstract This article constitutes the final and main part of a three-paper sequence (Ann PDE, 2016. doi:10.1007/s40818-016-0006-4; Oh and Tataru, 2015. arXiv:1503.01561), whose goal is to prove global well-posedness and scattering of the energy critical Maxwell-Klein-Gordon equation (MKG) on $\mathbb{R}^{1+4}$ for arbitrary finite energy initial data. Using the successively stronger continuation/scattering criteria established in the previous two papers (Ann PDE, 2016. doi:10.1007/s40818-016-0006-4; Oh and Tataru, 2015. arXiv:1503.01561), we carry out a blow-up analysis and deduce that the failure of global well-posedness and scattering implies the existence of a nontrivial stationary or self-similar solution to MKG. Then, by establishing that such solutions do not exist, we complete the proof.

Contents

✉ Daniel Tataru
tataru@math.berkeley.edu

Sung-Jin Oh
sjoh@math.berkeley.edu

1   Department of Mathematics, UC Berkeley, Berkeley, CA 94720, USA

1 Introduction

In this article we prove global well-posedness and scattering of the energy critical Maxwell-Klein-Gordon equation on $\mathbb{R}^{1+4}$ for any finite energy initial data. In Sect. 1.1, we present some background material concerning the Maxwell-Klein-Gordon equation on $\mathbb{R}^{1+4}$. Readers already familiar with this equation may skip to Sect. 1.2, where we give a precise statement of the main theorem (Theorem 1.3). This paper is the main and logically the final part of the three-paper sequence [26,27]. In Sects. 2 and 3 below, we provide an overview of the entire proof of Theorem 1.3 spanning the whole sequence.

1.1 (4 + 1)-dimensional Maxwell-Klein-Gordon system

Let $\mathbb{R}^{1+4}$ be the $(4 + 1)$-dimensional Minkowski space with the metric

$$\mathbf{m}_{\mu\nu} := \mathrm{diag}\,(-1, +1, +1, +1, +1)$$

in the standard rectilinear coordinates $(t = x^0, x^1, \ldots, x^4)$. Consider the trivial complex line bundle $L = \mathbb{R}^{1+4} \times \mathbb{C}$ over $\mathbb{R}^{1+4}$ with structure group $\mathrm{U}(1) = \{e^{i\chi} \in \mathbb{C}\}$. Global sections of L may be identified with $\mathbb{C}$-valued functions on $\mathbb{R}^{1+4}$. Using the identification $u(1) \equiv i\mathbb{R}$ and taking the trivial connection d as a reference, any connection $\mathbf{D}$ on L takes the form

$$\mathbf{D} = \mathrm{d} + i A$$

for some real-valued 1-form A on $\mathbb{R}^{1+4}$. The Maxwell-Klein-Gordon system is a Lagrangian field theory for a pair (A, ϕ) of a connection on L and a section of L with the action functional

$$\mathcal{S}[A, \phi] = \int_{\mathbb{R}^{1+4}} \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} \mathbf{D}_\mu \phi \overline{\mathbf{D}^\mu \phi}\, \mathrm{d}t\, \mathrm{d}x,$$

where $F_{\mu\nu} = (\mathrm{d}A)_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the curvature 2-form associated to $\mathbf{D}$. We follow the usual convention of raising/lowering indices by the Minkowski metric $\mathbf{m}$, and also of summing over repeated upper and lower indices. Computing the Euler-Lagrange equations, we arrive at the Maxwell-Klein-Gordon equations (MKG)

$$\begin{cases} \partial^\mu F_{\nu\mu} = \mathrm{Im}(\phi\overline{\mathbf{D}_\nu\phi}) \\ \Box_A\phi = 0, \end{cases} \tag{MKG}$$

where $\Box_A := \mathbf{D}^\mu\mathbf{D}_\mu$ is the (gauge) covariant d'Alembertian.

A basic feature of (MKG) is gauge invariance. Geometrically, a gauge transform is a change of basis in the fiber $\mathbb{C}$ over each point on $\mathbb{R}^{1+4}$ by an element of the gauge group U(1). Accordingly, we refer to a real-valued function $\chi : \mathbb{R}^{1+4} \to \mathbb{R}$ (hence $e^{i\chi} \in$ U(1)) as a gauge transformation and define the corresponding gauge transform of a pair (A, ϕ) as

$$(A, \phi) \mapsto (\widetilde{A}, \widetilde{\phi}) := (A - \mathrm{d}\chi, e^{i\chi}\phi). \tag{1.1}$$

Observe that $\mathbf{D}$ and $\Box_A$ are covariant under gauge transforms (i.e., $e^{i\chi}\mathbf{D}\phi = \widetilde{\mathbf{D}}\widetilde{\phi}$ etc), whereas F and $\mathrm{Im}(\phi\overline{\mathbf{D}_\mu\phi})$ are invariant. Hence (MKG) is invariant under gauge transforms. Since U(1) is an abelian group, (MKG) is said to be an abelian gauge theory.

We now formulate the initial value problem for (MKG), in a way that is consistent with the gauge invariance of the system. An initial data set for (MKG) consists of a pair of 1-forms (a_j, e_j) and a pair of $\mathbb{C}$-valued functions (f, g) on $\mathbb{R}^4$. We say that (a, e, f, g) is the initial data for a solution (A, ϕ) at time t_0 if

$$(A_j, F_{0j}, \phi, \mathbf{D}_t\phi)\restriction_{\{t=t_0\}} = (a_j, e_j, f, g).$$

We usually take the initial time t_0 to be zero. Observe that the $\nu = 0$ component of (MKG) imposes a constraint on any initial data for (MKG), namely

$$\partial^\ell e_\ell = \mathrm{Im}(f\overline{g}) \tag{1.2}$$

This equation is called the Gauss (or constraint) equation.

There is a conserved energy for (MKG), which is one of the basic ingredients of the non-perturbative analysis performed in this paper. We define the conserved energy of a solution (A, ϕ) at time t to be

$$\mathcal{E}_{\{t\}\times\mathbb{R}^4}[A, \phi] := \frac{1}{2}\int_{\{t\}\times\mathbb{R}^4} \sum_{0\le\mu<\nu\le4} |F_{\mu\nu}|^2 + \sum_{0\le\mu\le4} |\mathbf{D}_\mu\phi|^2 \, \mathrm{d}x. \tag{1.3}$$

For a suitably regular solution to (MKG) defined on a connected interval I, this quantity is constant. This conservation law is in fact a consequence of Nöther's principle (i.e., continuous symmetry of the field theory corresponds to a conserved quantity) applied to the time translation symmetry of (MKG); we refer to Sect. 5 for further discussion and a proof.

Observe that the conserved energy is invariant under the scaling

$$(A, \phi)(t, x) \mapsto \left(\lambda^{-1} A, \lambda^{-1} \phi\right) \left(\lambda^{-1} t, \lambda^{-1} x\right) \quad \text{for any } \lambda > 0,$$

which also preserves the system (MKG). Hence (MKG) on $\mathbb{R}^{1+4}$ is energy critical.

1.2 Statement of the main theorem

Our goal now is to give a precise statement of the global well-posedness/scattering theorem proved in this paper. For this purpose, we first borrow some definitions from [19,26].

We say that a (MKG) initial data set (a, e, f, g) (i.e., a solution to the Gauss equation) is classical and write $(a, e, f, g) \in \mathcal{H}^{\infty}$ if each of a, e, f, g belongs to $H_x^{\infty} := \cap_{n=0}^{\infty} H_x^n$. Correspondingly, we say that a smooth solution (A, ϕ) to (MKG) on $I \times \mathbb{R}^4$ (where $I \subseteq \mathbb{R}$ is an interval) is a classical solution if $A_{\mu}, \phi \in \cap_{n,m=0}^{\infty} C_t^m(I; H_x^n)$.

Define the space $\mathcal{H}^1 = \mathcal{H}^1(\mathbb{R}^4)$ of finite energy initial data sets to be the space of (MKG) initial data sets for which the following norm is finite:

$$\|(a, e, f, g)\|_{\mathcal{H}^1} := \sup_{j=1,\dots,4} \|(a_j, e_j)\|_{\dot{H}_x^1 \times L_x^2(\mathbb{R}^4)} + \|(f, g)\|_{\dot{H}_x^1 \times L_x^2(\mathbb{R}^4)}.$$
$$(1.4)$$

Given a pair (A, ϕ) on $I \times \mathbb{R}^4$, we define its $C_t \mathcal{H}^1(I \times \mathbb{R}^4)$ norm as

$$\|(A, \phi)\|_{C_t \mathcal{H}^1(I \times \mathbb{R}^4)} := \operatorname{ess\,sup}_{t \in I} \left(\|A[t]\|_{\dot{H}_x^1 \times L_x^2} + \|\phi[t]\|_{\dot{H}_x^1 \times L_x^2} \right),$$

where $A[t]$ and $\phi[t]$ are shorthands for $(A, \partial_t A)(t)$ and $(\phi, \partial_t \phi)(t)$, respectively. We then define the notion of an admissible $C_t \mathcal{H}^1$ solution to (MKG) via approximation by classical solutions as follows.

Definition 1.1 (*Admissible $C_t \mathcal{H}^1$ solutions to* (MKG)) Let $I \subseteq \mathbb{R}$ be an interval. We say that a pair $(A, \phi) \in C_t \mathcal{H}^1(I \times \mathbb{R}^4)$ is an admissible $C_t \mathcal{H}^1(I \times \mathbb{R}^4)$ solution to (MKG) if there exists a sequence $(A^{(n)}, \phi^{(n)})$ of classical solutions to (MKG) on $I \times \mathbb{R}^4$ such that

$$\|(A, \phi) - (A^{(n)}, \phi^{(n)})\|_{C_t \mathcal{H}^1(J \times \mathbb{R}^4)} \to 0 \quad \text{as } n \to \infty,$$

for every compact subinterval $J \subseteq I$.

The necessity of restricting the class of energy solutions under consideration to the admissible ones as defined above is a relatively standard matter in the

realm of low regularity solutions for nonlinear dispersive equations. Often uniqueness statements require additional regularity properties for solutions, which are then proved to hold for the solutions which are limits of smooth solutions, but might not be true or straightforward in general. In our case the difficulties are compounded by the need to have a good notion of finite energy solution which is gauge invariant.

Remark 1.2 The above definitions can be localized to an open subset $O \subseteq \mathbb{R}^4$ or $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ in an obvious manner; see [26, Sects. 3 and 5].

Next, we recall the global Coulomb gauge condition

$$\partial^\ell A_\ell = \sum_{\ell=1,\dots,4} \partial_\ell A_\ell = 0. \tag{1.5}$$

The role of this condition is to fix the ambiguity arising from the gauge invariance of (MKG), which is an immediate formal obstruction for well-posedness.

Finally, given an interval $I \subseteq \mathbb{R}$, we borrow the space-time norms $Y^1(I \times \mathbb{R}^4)$ and $S^1(I \times \mathbb{R}^4)$ from [19,26,27]. We define the S^1 norm of a solution (A, ϕ) on $I \times \mathbb{R}^4$ to be

$$\|(A, \phi)\|_{S[I]} := \|A_0\|_{Y^1(I \times \mathbb{R}^4)} + \|A_x\|_{S^1(I \times \mathbb{R}^4)} + \|\phi\|_{S^1(I \times \mathbb{R}^4)}.$$

In particular, the S^1 norm captures the dispersive properties of A_x and ϕ. The precise definition of the S^1 norm is rather intricate; instead of the full definition, in this paper we only rely on a few basic properties of the spaces Y^1 and S^1, such as those below (see also Remark 4.2).

$$\|(\varphi, \partial_t \varphi)\|_{C_t(I; \dot{H}_x^1 \times L_x^2)} \lesssim \|\varphi\|_{S^1(I \times \mathbb{R}^4)},$$
$$\|(\varphi, \partial_t \varphi)\|_{C_t(I; \dot{H}_x^1 \times L_x^2)} \lesssim \|\varphi\|_{Y^1(I \times \mathbb{R}^4)}.$$

We are now ready to state our main theorem.

Theorem 1.3 (Main Theorem) *Let $(a, e, f, g) \in \mathcal{H}^1$ be a finite energy initial data set for* (MKG) *obeying the global Coulomb gauge condition $\partial^\ell a_\ell = 0$. Then there exists a unique admissible $C_t \mathcal{H}^1$ solution (A, ϕ) to the initial value problem defined on the whole $\mathbb{R}^{1+4}$ which satisfies the global Coulomb gauge condition $\partial^\ell A_\ell = 0$. Moreover, the S^1 norm of (A, ϕ) is finite, i.e.,*

$$\|A_0\|_{Y^1(\mathbb{R}^{1+4})} + \|A_x\|_{S^1(\mathbb{R}^{1+4})} + \|\phi\|_{S^1(\mathbb{R}^{1+4})} < \infty. \tag{1.6}$$

Remark 1.4 The a-priori bound above implies scattering towards both $t \to \pm\infty$; see Theorem 4.8. It also implies continuity of the data to solution map on compact time intervals, though not on the full real line.

Remark 1.5 We do not lose any generality by restricting to initial data sets in the global Coulomb gauge, since any finite energy initial data set can be gauge transformed to obey the condition $\partial^\ell a_\ell = 0$. See [26, Sect. 3].

Remark 1.6 We note that an independent proof of global well-posedness and scattering of MKG-CG has been recently given by Krieger-Lührmann [16], following a version of the Bahouri-Gérard nonlinear profile decomposition [1] and Kenig-Merle concentration compactness/rigidity scheme [8,9] developed by Krieger-Schlag [17] for the energy critical wave maps. We refer to Sect. 3.2 for a brief comparison between our work and [16].

1.3 A brief history and broader context

A natural point of view is to place the present papers and results within the larger context of nonlinear wave equations, of which the starting point is the semilinear wave equation $\Box u = \pm |u|^p u$. More accurately, the (MKG) equation belongs to the class of geometric wave equations, which includes wave maps (WM), Yang-Mills (YM), Einstein equations, as well as many other coupled models. Two common features of all these problems are that they admit a Lagrangian formulation, and have some natural gauge invariance properties. Following are some of the key developments that led to the present work.

1. The null condition A crucial early observation in the study of both long range and low regularity solutions to geometric wave equations was that the nonlinearities appearing in the equations have a favorable algebraic structure, which was called null condition, and which can be roughly described as a cancellation condition in the interaction of parallel waves. In the low regularity setting, this was first explored in work of Klainerman and Machedon [10], and by many others later on.

2. The $X^{s,b}$ spaces A second advance was the introduction of the $X^{s,b}$ spaces,[1] also first used by Klainerman and Machedon [13] in the context of the wave equation. Their role was to provide enough structure in order to be able to take advantage of the null condition in bilinear and multilinear estimates. Earlier methods, based on energy bounds, followed by the more robust Strichartz estimates, had proved inadequate to the task.

3. The null frame spaces To study nonlinear problems at critical regularity one needs to work in a scale invariant setting. However, it was soon realized that the homogeneous $X^{s,b}$ spaces are not even well defined, not to mention suitable for this. The remedy, first introduced in work of the second author [41] in the context of wave maps, was to produce a better description of the fine

---

[1] The concept, and also the notation, is due to Bourgain, in the context of KdV and NLS type problems.

structure of waves, combining frequency and modulation localizations with adapted frames in the physical space. This led to the null frame spaces, which played a key role in subsequent developments for wave maps. We remark that another scale invariant alternative to $X^{s,b}$ spaces are the U^p and V^p spaces, also originally developed by the second author; while these played a role in the study of other nonlinear dispersive problems at critical regularity, they play no role in the present story.

4. Renormalization A remarkable feature of all semilinear geometric wave equations is that while at high regularity (and locally in time) the nonlinearity is perturbative, this is no longer the case at critical regularity. Precisely, isolating the non-perturbative component of the nonlinearity, one can see that this is of paradifferential type; in other words, the high frequency waves evolve on a variable low frequency background. To address this difficulty, the idea of Tao [34], also in the wave map context, was to renormalize the paradifferential problem, i.e., to find a suitable approximate conjugation to the corresponding constant coefficient problem.

5. Induction of energy The ideas discussed so far seem to suffice for small data critical problems. Attacking the large data problem generates yet another range of difficulties. One first step in this direction is Bourgain's induction of energy idea [2], which is a convenient mechanism to transfer information to higher and higher energies. We remark that an alternate venue here, which sometimes yields more efficient proofs, is the Kenig-Merle idea [9] of constructing minimal blow-up solutions. However, the implementation of this method in problems which require renormalization seems to cause considerable trouble.

For a further discussion on this issue, we refer to the work of Krieger-Schlag [17], where this method was carried out in the case of energy critical wave maps into the hyperbolic plane. We also mention the recent paper [16], where an independent proof of global well-posedness and scattering for (MKG) in the Coulomb gauge was given following the above-mentioned ideas of Kenig-Merle and Krieger-Schlag. See Sect. 3.2 for a brief comparison between the approaches in this paper and [16].

6. Energy dispersion One fundamental goal in the study of large data problems is to establish a quantitative dichotomy between dispersion and concentration. The notion of energy dispersion, introduced in joint work [32,33] of the second author and Sterbenz in the wave map context, provides a convenient measure for pointwise concentration. Precisely, at each energy there is an energy dispersion threshold below which dispersion wins. We remark that, when it can be applied, the Kenig-Merle method [9] yields more accurate information; for instance, see [17]. However, the energy dispersion idea, which is what we follow in the present series of papers, is much easier to implement in conjunction with renormalization.

7. The frequency gap One obstacle in the transition from small to large data in renormalizable problems is that the low frequency background may well correspond to a large solution. Is this fatal to the renormalized solution? The answer to that, also originating in [32,33], is that there may be a second hidden source of smallness, namely a large frequency gap between the high frequency wave and the low frequency background it evolves on.

8. Morawetz estimates The outcome of the ideas above is a dichotomy between dispersion and scattering on one hand, and very specific concentration patterns, e.g., solitons, self-similar solutions on the other hand. The Morawetz estimates, first appearing in this role in the work of Grillakis [6], are a convenient and relatively simple tool to eliminate such concentration scenarios.

We now recall some earlier developments on geometric wave equations related to the present paper. We start our discussion with the (MKG) problem above the scaling critical regularity. In the two and three dimensional cases, which are energy subcritical, global regularity of sufficiently regular solutions was shown in the early works [4,5,23]. The former two in fact handled the more general Yang-Mills-Higgs system. In dimension $d = 3$, this result was greatly improved by [11], which established global well-posedness for any finite energy data. In this work, the quadratic null structure of (MKG) in the Coulomb gauge was uncovered and used for the first time. Subsequent developments were made by [3] and more recently [22], where an essentially optimal local well-posedness result was established. An important observation in [22] is that (MKG) in Coulomb gauge exhibits a secondary multilinear cancellation feature. The related paper [7] is concerned with global well-posedness of the same problem at low regularity. We also mention the work [30], in which finite energy global well-posedness was established in the Lorenz gauge. In the higher dimensional case $d \geq 4$, an essentially optimal local well-posedness result for a model problem closely related to (MKG) was obtained in [15]. This was followed by further refinements in [29,31].

The progress for the closely related Yang-Mills system (YM) in the subcritical regularity has largely paralleled that of (MKG), at least for small data. Indeed, (YM) exhibits a null structure in the Coulomb gauge which is very similar to (MKG). In particular, the aforementioned work [15] is also relevant for the small data problem for (YM) in the Coulomb gauge at an essentially optimal regularity.

However, a new difficulty arises in the large data[2] problem for (YM): namely, the gauge transformation law is nonlinear due to the non-abelian

---

[2] More precisely, a suitable scaling critical norm of the connection A (e.g., $\|A\|_{L_x^d}$) or the curvature F (e.g., $\|F\|_{L_x^{\frac{d}{2}}}$) is large.

gauge group. In particular, gauge transformations into the Coulomb gauge obey a nonlinear elliptic equation, for which no suitable large data regularity theory is available. Note, in comparison, that such gauge transformations obey a linear Poisson equation in the case of (MKG). In [12], where finite energy global well-posedness of the $3 + 1$ dimensional (YM) problem was proved, this issue was handled by localizing in space-time via the finite speed of propagation to gain smallness, and then working in local Coulomb gauges. An alternative, more robust approach without space-time localizations to the same problem has been put forth by the first author in [24,25], inspired by [36–40]. The idea is to use an associated geometric flow, namely the Yang-Mills heat flow, to select a global-in-space Coulomb-like gauge for data of any size.

Before turning to the (MKG) and (YM) problems at critical regularity, we briefly recall some recent developments on the wave map equation (WM), where many of the methods we implement here have their roots. We confine our discussion to the energy critical problem in $2 + 1$ dimensions, which is both the most difficult and the most relevant to our present paper. For the small data problem, global well-posedness was established in [34,35,42]. More recently, the threshold theorem for large data wave maps, which asserts that global well-posedness and scattering hold below the ground state energy, was proved in [32, 33] when the target $\mathcal{N}$ is a general compact manifolds. Independently, global well-posedness and scattering for (WM) were established in [17] when $\mathcal{N}$ is the hyperbolic plane and in [36–40] when $\mathcal{N}$ is a general hyperbolic space. In both cases, $\mathcal{N}$ is a non-compact manifold which do not admit any nontrivial finite energy harmonic maps from $\mathbb{R}^2$; moreover, an a-priori bound of the scattering norm in terms of the energy followed immediately from the proofs of [17,36–40]. In fact, [17] also established a nonlinear profile decomposition for bounded (in energy) sequences of wave maps. See also [20] for a sharp refinement in the case of a two-dimensional target, taking into account an additional topological invariant (namely, the degree of the wave map). Our present strategy was strongly influenced by [32,33], which can be seen as the first predecessor of this work.

Despite the many similarities, there is a key structural difference between (WM) on the one hand and (MKG), (YM) on the other, whose understanding is crucial for making progress on the latter two problems. Roughly speaking, all three equations can be written in a form where the main 'dynamic variables', which we denote by ϕ, obey a possibly nonlinear gauge covariant wave equation $\Box_A \phi = \cdots$, and the associated curvature $F[A]$ is determined by ϕ. In the case of (WM), this dependence is simply algebraic, whereas for (MKG) and (YM) the curvature $F[A]$ obeys a wave equation with a nonlinearity depending on ϕ. This difference manifests in the renormalization procedure for each equation: For (WM) it suffices to use a physical space gauge transformation,

whereas for (MKG) and (YM) it is necessary to use a microlocal (more precisely, pseudo-differential) gauge transformation that exploits the fact that A solves a wave equation in a suitable gauge.

The first (MKG) renormalization argument appeared in [28], in which global regularity of (MKG) for small critical Sobolev data was established in dimensions $d \geq 6$. This work was followed by a similar high dimensional result for (YM) in [18]. Finally, the small data result in the energy critical dimension $4 + 1$ was obtained in [19], which may be viewed as the second direct predecessor to the present work. In particular we borrow a good deal of notations, ideas and estimates from [19].

We end our introduction with a few remarks on the energy critical (YM) problem in $4 + 1$ dimensions, which is a natural next step after the present work. The issue of non-abelian gauge group for the large data problem has already been discussed. Another important difference between (MKG) and (YM) in $4 + 1$ dimensions is that the latter problem admits instantons, which are nontrivial static solutions with finite energy. Therefore, in analogy with (WM), it is reasonable to put forth the threshold conjecture for the energy critical (YM) problem, namely that global well-posedness and scattering hold below the energy of the first instanton. Finally, (YM) is more 'strongly coupled' as a system compared to (MKG), in the sense that the connection A itself obeys a covariant wave equation. This feature seems to necessitate a more involved renormalization procedure compared to (MKG).

2 Overview of the proof I: summary of the first two papers

The basic strategy for proving Theorem 1.3 is by contradiction, following the scheme successfully developed in [32,33] in the setting of energy critical wave maps. In the first two papers of the sequence [26,27] we establish successively stronger continuation and scattering criteria, whose contrapositives provide precise information about the nature of a finite time blow-up (i.e., failure of global well-posedness) or non-scattering. In the present paper, we use this information, as well as conservation laws and Morawetz-type monotonicity formulae for (MKG), to perform a blow-up analysis and show that the failure of Theorem 1.3 implies the existence of a nontrivial finite energy stationary or self-similar solution to (MKG). Since such a solution does not exist (see Sect. 7 below), Theorem 1.3 must hold.

In this section we review the main results and ideas of the earlier two papers in the sequence [26,27]. In Sect. 3 we summarize the argument given in the present paper. To steer away from unnecessary technical details we only consider smooth data and solutions; however we remark that the results also apply

to merely finite energy data and admissible $C_t \mathcal{H}^1$ solutions. For the notation, we refer to Sect. 4.

2.1 Local well-posedness in the global Coulomb gauge and non-concentration of energy

The main result of the first paper [26] of the sequence is local well-posedness of (MKG) in the global Coulomb gauge with a lower bound on the lifespan in terms of the energy concentration scale

$$r_c = r_c(E)[a, e, f, g] := \sup \left\{ r > 0 : \forall x \in \mathbb{R}^4, \mathcal{E}_{B_r(x)}[a, e, f, g] < \delta_0(E, \epsilon_*^2) \right\},$$

where $B_r(x)$ denotes the open ball of radius r with center x, $\delta_0(E, \epsilon_*^2) = c^2 \epsilon_*^2 \min \left\{ 1, \epsilon_*^4 E^{-2} \right\}$, c is an absolute constant and ϵ_*^2 is the threshold for the small energy global well-posedness result in [19] (see Theorem 4.1). A simplified version of the main theorem of [26] is as follows:

Theorem 2.1 *Given any $E > 0$ let $\delta_0(E, \epsilon_*^2) > 0$ be as above. Let (a, e, f, g) be a smooth finite energy initial data for (MKG) satisfying the global Coulomb gauge condition $\sum_j \partial_j a_j = 0$. Then there exists a unique smooth solution (A, ϕ) to (MKG) in the global Coulomb gauge on $[-r_c, r_c] \times \mathbb{R}^4$.*

Theorem 2.1 implies that finite time blow-up is always accompanied by concentration of energy (i.e., $r_c \to 0$ at the end of the maximal lifespan). For a precise statement, see Theorem 4.3. In what follows we explain the ideas involved in the proof of local existence, which lies at the heart of Theorem 2.1.

2.1.1 Strategy of proof in model cases

For many other semi-linear equations, such as $\Box u = \pm u^{\frac{d+2}{d-2}}$ or the wave map equation, a result analogous to Theorem 2.1 is a rather immediate consequence of small energy global well-posedness and finite speed of propagation. Roughly speaking, the proof (of local existence) proceeds in the following three steps:

> *Step A.* One truncates the initial data locally in space to achieve small energy.
> *Step B.* By small energy global well-posedness, the truncated data give rise to global solutions. Restricting these global solutions to the domain of dependence of the truncated regions, one obtains a family of local-in-spacetime solutions that agree with each other on the intersection of their domains by finite speed of propagation.

Step C. One patches together these solutions to obtain a local-in-time solution to the original initial data.

In particular, the lifespan of the solution constructed by this scheme depends on the size of spatial truncation in Step A, which in turn is dictated by the energy concentration scale r_c of the initial data.

2.1.2 Non-locality of (MKG) in the global Coulomb gauge

When carrying out the above strategy in our setting, however, we face difficulties arising from non-local features of (MKG) in the global Coulomb gauge. One source of non-locality is the Gauss (or the constraint) equation

$$\partial^\ell e_\ell = \text{Im}(f\overline{g}), \tag{2.1}$$

which must be satisfied by every (MKG) initial data set. Another source is the presence of the elliptic equation for A_0 in the global Coulomb gauge (cf. (2.4)); in particular, finite speed of propagation fails in the global Coulomb gauge.

In the remainder of this subsection, we give an overview of the techniques developed in [26] for overcoming these issues, and explain how these can be used to essentially execute Steps A-C above to obtain Theorem 2.1 from the small energy global well-posedness theorem proved in [19] (see Theorem 4.1).

2.1.3 Execution of step A: initial data excision and gluing

Consider the problem of truncating a (MKG) initial data set[3] (a, e, f, g) to a ball B. A naive way to proceed would be to apply a smooth cutoff to each of a, e, f, g. However, integrating the Gauss equation (2.1) by parts over balls of large radius, we see that e_j must in general be nontrivial on the boundary spheres outside B, even if f and g are supported in B.

Instead, the idea of initial data excision and gluing[4] is as follows: rather than just excising the unwanted part, we glue it to another initial data set (i.e., solution to the Gauss equation) which has an explicit description, so that the

---

[3] In application a obeys the global Coulomb gauge condition $\partial^\ell a_\ell = 0$, but this fact is irrelevant for the discussion here.

[4] We remark that similar techniques have been developed in mathematical general relativity, as a means to construct a large class of interesting initial data sets for the Einstein equations. Our setting involves a simpler constraint equation, but we require sharp techniques which are applicable at the critical regularity.

Gauss equation is still satisfied. For example, in the exterior of a ball B we may glue to the data

$$\left(e_{(q)j} = \frac{q}{2\pi^2} \frac{x_j}{|x|^4}, 0, 0, 0 \right)$$

with an appropriate q. Note that $e_{(q)}$ is precisely the electric field of an electric monopole of charge q placed at the origin.

Using this idea we may truncate (a, e, f, g) to balls to make the energy sufficiently small. The minimum size of these balls, which later dictates the lifespan of the solution, can be chosen to be proportional to the energy concentration scale. This procedure is our analogue of step A.

2.1.4 Execution of step B: geometric uniqueness of admissible solution to (MKG)

Though finite speed of propagation fails for (MKG) in certain gauges such as the global Coulomb gauge, it is still true up to gauge transformations. We refer to this statement as local geometric uniqueness for (MKG), and use it as a substitute for the usual finite speed of propagation property.

Applying a suitable gauge transformation to each truncated initial data set to impose the global Coulomb gauge condition, we are in position to apply the small energy global well-posedness theorem (Theorem 4.1) and construct a family of global smooth solutions. Restricting these solutions to the domain of dependence of the truncated regions and appealing to local geometric uniqueness, we obtain local-in-spacetime Coulomb solutions (i.e., obey $\partial^\ell A_\ell = 0$ on the domains) which are gauge equivalent to each other on the intersection of their domains. We refer to such solutions as compatible pairs[5]; geometrically, these are precisely local descriptions of a globally defined pair of a connection and a section on local trivializations of the bundle L.

2.1.5 Execution of step C: patching local Coulomb solutions

The final task is to patch together the local-in-spacetime descriptions of a solution (i.e., compatible pairs) to produce a global-in-space solution (A, ϕ) in the global Coulomb gauge. We first adapt a patching argument of Uhlenbeck [43] to produce a single global-in-space solution (A', ϕ') obeying an appropriate S^1 norm bound. The fact that a gauge transformation χ between Coulomb gauges obeys the Laplace equation $\triangle \chi = 0$, and hence possesses improved regularity, is important for this step. The solution (A', ϕ') obtained by this

---

[5] See also Sect. 6.3 of the present paper, where this notion arises naturally from local limits of a sequence of solutions.

patching process is not necessarily in the global Coulomb gauge; it is however approximately Coulomb (i.e., $\partial^\ell A'_\ell$ obeys an improved bound), since it arose from patching together local Coulomb solutions. It is thus possible to find a nicely behaved gauge transformation into the exact global Coulomb gauge, leading us to the desired local-in-time solution.

2.2 Continuation of energy dispersed solutions

We now describe the content of [27]. The main theorem of [27] is a continuation/scattering criterion in the global Coulomb gauge for a large energy solution (A, ϕ) to (MKG) in terms of its energy dispersion $ED[\phi](I)$, defined as

$$ED[\phi](I) = \sup_k \left(2^{-k} \|P_k \phi\|_{L^\infty_{t,x}(I \times \mathbb{R}^4)} + 2^{-2k} \|\partial_t P_k \phi\|_{L^\infty_{t,x}(I \times \mathbb{R}^4)} \right) \quad (2.2)$$

for any time interval $I \subseteq \mathbb{R}$. A simple version is as follows:

Theorem 2.2 *Given any $E > 0$, there exist positive numbers $\epsilon = \epsilon(E) > 0$ and $F = F(E)$ such that the following holds. Let (A, ϕ) be a smooth solution to* (MKG) *in the global Coulomb gauge (MKG-CG) on $I \times \mathbb{R}^4$ with energy $\leq E$. If $ED[\phi](I) \leq \epsilon(E)$, then the following a-priori S^1 norm bound holds:*

$$\|A_0\|_{Y^1[I]} + \|A_x\|_{S^1[I]} + \|\phi\|_{S^1[I]} \leq F(E). \quad (2.3)$$

Moreover, (A, ϕ) extends as a smooth solution past finite endpoints of I.

Theorem 2.2 is analogous to the main result in [32] for energy critical wave maps. Thanks to the a-priori bound (2.3), the solution (A, ϕ) scatters towards each infinite endpoint in the sense of Remark 1.4. For a more precise formulation, see Theorems 4.7 and 4.8.

We now describe the main ideas of the proof of Theorem 2.2. In what follows, we only consider solutions to (MKG) in the global Coulomb gauge.

2.2.1 Decomposition of the nonlinearity

We begin by describing the structure of the Maxwell-Klein-Gordon system in the global Coulomb gauge (MKG-CG), which take the form

$$\begin{cases} \triangle A_0 = \mathrm{Im}(\phi \overline{\partial_t \phi}) + \text{(cubic terms)} \\ \Box A_j = \mathcal{P}_j \mathrm{Im}(\phi \overline{\partial_x \phi}) + \text{(cubic terms)} \\ \Box \phi = -2i A_\mu \partial^\mu \phi + \text{(cubic terms)} \end{cases} \quad (2.4)$$

where $\mathcal{P}$ is the Leray L^2-projection to the space of divergence-free vector fields. We omitted cubic terms as they are strictly easier to handle. The elliptic equation for A_0 allows us to obtain the appropriate Y^1 bound once we establish S^1 bounds for A_x and ϕ; henceforth we focus on the wave equations for A_x and ϕ.

As in the case of small energy global well-posedness [19], the null structure of (MKG) in the global Coulomb gauge plays an essential role in the proof of Theorem 2.2. All quadratic terms in the wave equations exhibit null structure, i.e., cancellation in the angle between inputs in Fourier space. There is also a secondary multilinear null structure in the term $2i A_\mu \partial^\mu \phi$ which arises by plugging in the equations for A_0, A_j. All of this structure is necessary for controlling the S^1 norm of (A, ϕ), but it is by no means sufficient as we discuss below.

2.2.2 Renormalization for large energy

Even in the case of small energy global well-posedness [19], the null structure alone is not enough to bound the S^1 norm of (A, ϕ) due to the paradifferential term in the ϕ-equation

$$- \sum_k 2i\, P_{<k} A^{\text{free}} \cdot \partial_x P_k \phi.$$

Here A_j^{free} is the free wave evolution of $A_j[0] := (A_j, \partial_t A_j) \upharpoonright_{\{t=0\}}$. As in [19,28], we handle this term by a renormalization argument. More precisely, we treat the problematic term as a part of the linear operator and construct a paradifferential parametrix. The construction in [19,28], however, relied on smallness of the energy, which we lack in our setting. Instead we consider the linear operator with a frequency gap m

$$\Box_{A^{\text{free}}}^{p,m} \psi := \Box \psi + \sum_k 2i\, P_{<k-m} A_x^{\text{free}} \cdot \partial_x P_k \psi,$$

and gain smallness by taking m sufficiently large. This idea is akin to the gauge renormalization procedure for wave maps in [32], where a large frequency gap was used to control the large paradifferential term.

2.2.3 Role of energy dispersion

We now describe the role of small energy dispersion $ED[\phi]$. Roughly speaking, small energy dispersion allows us to gain in transversal balanced frequency interactions. This complements the gain in parallel interactions, due to the null condition, and the gain in the $high \times high \rightarrow low$ interactions due to the

favorable frequency balance. For instance, by interpolation with (non-sharp) Strichartz norms controlled by the S^1 norm, we have[6]

$$\|P_k(P_{k_1}\phi P_{k_2}\psi)\|_{L^2_{t,x}(I\times\mathbb{R}^4)} \lesssim 2^{-\frac{1}{2}\min\{k_1,k_2\}} ED[\phi]^\theta \|P_{k_1}\phi\|_{S^1[I]}^{1-\theta} \|P_{k_2}\psi\|_{S^1[I]},$$
(2.5)

which is useful when $k_1 = k + O(1)$, $k_2 = k + O(1)$ and ϕ, ψ are at a large angle so that the output modulation is high.

To see how this gain is useful, we return to the full nonlinear system (MKG) in the global Coulomb gauge. Upon decomposing the inputs and output into Littlewood-Paley pieces, most of the nonlinearity exhibits an off-diagonal exponential decay in frequency. For example, the nonlinearity in the A_x-equation obeys

$$\|P_k \mathcal{P}_x(P_{k_1}\phi \partial_x \overline{P_{k_2}\phi})\|_{N[I]} \lesssim 2^{-\delta(|k-k_1|+|k-k_2|)} \|P_{k_1}\phi\|_{S^1[I]} \|P_{k_2}\phi\|_{S^1[I]}.$$

Introducing again a large frequency gap m, we gain smallness except when $k_1 = k + O_m(1)$ and $k_2 = k + O_m(1)$. Furthermore, thanks to the null structure, we also gain extra smallness except for angled interaction; then we are precisely in position to use $ED[\phi]$. In conclusion, we gain smallness from $ED[\phi] \leq \epsilon$ for the nonlinearity in the A_x-equation.

2.2.4 Linear well-posedness of $\Box_A \psi = f$

Unfortunately the a-priori estimate (2.3) does not close yet, as there exists a nonlinear term in the ϕ-equation with no off-diagonal exponential decay. This part is precisely the $low \times high \to high$ frequency and $high \times low \to low$ modulation interaction[7] in the term $-2iA \cdot \partial_x\phi$, i.e.,

$$-2i \sum_{\substack{k_1 < k \\ k_2 = k+O(1)}} \sum_{j < k_1} P_k Q_{<j}(P_{k_1} Q_j A \cdot \partial_x P_{k_2} Q_{<j}\phi).$$
(2.6)

Nevertheless, this term has the redeeming feature that it can be bounded by a *divisible* norm: Given any $\varepsilon > 0$ the interval I can be split into smaller pieces I_k on each of which the N norm of the above expression is bounded by $\leq \varepsilon^2 \|P_{k_2}\phi\|_{S^1[I]}$, where the number of such intervals is $O_{\|\phi\|_{S^1[I]},\varepsilon}(1)$. For a

---

[6] Note that (2.5) is symmetric in ϕ and ψ, so we may choose to use the energy dispersion norm of either. Note also that all nonlinearity of (MKG) involve at least one factor of ϕ. This is why it suffices to assume smallness of just $ED[\phi]$ and not A.

[7] We note that this term is where the secondary multilinear cancellation structure of MKG-CG is needed.

solution (A, ϕ) to (MKG), this observation leads to linear well-posedness of
the magnetic wave equation[8] $\Box_A \psi = f$ with bound

$$\|\psi\|_{S^1[I]} \lesssim_{\|(A_x, \phi)\|_{S^1[I]}} \|\psi[0]\|_{\dot{H}^1_x \times L^2_x} + \|f\|_{N[I]}, \tag{2.7}$$

where $\psi[0] := (\psi, \partial_t \psi)\upharpoonright_{\{t=0\}}$. The bound (2.7) allows us to setup an *induction
on energy* scheme to establish (2.3), which we now turn to explain.

2.2.5 Induction on energy

The starting point of our induction is the small energy global well-posedness
theorem [19], which implies that (2.3) holds with $F(E) = C\sqrt{E}$ when the
energy E is sufficiently small. Our goal is to show the existence of a non-
increasing positive function $c_0(\cdot)$ on the whole interval $[0, \infty)$ such that if the
conclusion of Theorem 2.2 holds for energy up to E, then it also holds for
energy up to $E + c_0(E)$. Monotonicity of $c_0(\cdot)$ implies that it has a uniform
positive lower bound on every finite interval; thus the continuous induction
works for all energy.

In what follows, we describe the construction of $c_0(E)$, $F := F(E + c_0(E))$
and $\epsilon := \epsilon(E + c_0(E))$ under the induction hypothesis that Theorem 2.2 holds
up to energy E for some $F(E)$ and $\epsilon(E)$. For the scheme to work, it is crucial
to let $c_0(E)$ depend only on E and not on $F(E)$ or $\epsilon(E)$. On the other hand,
F and ϵ may depend on $F(E)$ and $\epsilon(E)$.

Let (A, ϕ) be a solution on $I \times \mathbb{R}^4$ with energy $E + c_0(E)$ and $ED[\phi] \leq \epsilon$.
To prove (2.3) for (A, ϕ), we compare it with another solution $(\tilde{A}, \tilde{\phi})$ with
frequency truncated initial data[9]

$$(\tilde{A}_j[0], \tilde{\phi}[0]) = (P_{\leq k^*} A_j[0], P_{\leq k^*} \phi[0])$$

where the 'cut frequency' $k^* \in \mathbb{R}$ is chosen so that $(\tilde{A}, \tilde{\phi})$ has energy E. By
taking $c_0(E)$ and ϵ sufficiently small, we aim for the following two goals:

Goal A. The energy dispersion $ED[\tilde{\phi}](I)$ is sufficiently small so that the
induction hypothesis applies to $(\tilde{A}, \tilde{\phi})$. Hence

$$\|\tilde{A}_0\|_{Y^1[I]} + \|\tilde{A}_x\|_{S^1[I]} + \|\tilde{\phi}\|_{S^1[I]} \leq F(E). \tag{2.8}$$

---

[8] More precisely, the observation regarding (2.6), combined with the paradifferential para-
metrix construction mentioned above, implies well-posedness of the equation $\Box_A^{p,m} \psi :=$
$\Box\psi + 2i \sum_k P_{<k-m} A_\mu \partial^\mu P_k \psi = f$ for sufficiently large m with bound (2.7). The terms in
$\Box_A - \Box_A^{p,m}$ also turn out to be bounded by divisible norms, which leads to the well-posedness
of $\Box_A \psi = f$.

[9] In the global Coulomb gauge, $A_x[0] = (A_x, \partial_t A_x)(0)$ and $\phi[0] = (\phi, \partial_t \phi)(0)$ determine the
whole initial data set (a, e, f, g), as we can solve for A_0 in the constraint equation $-\triangle A_0 =$
$\text{Im}(\phi \overline{\partial_t \phi}) - |\phi|^2 A_0$.

Goal B. The difference $(B^{high}, \psi^{high}) := (A_\mu - \tilde{A}_\mu, \phi - \tilde{\phi})$ obeys

$$\|B_0^{high}\|_{Y^1[I]} + \|B_x^{high}\|_{S^1[I]} + \|\psi^{high}\|_{S^1[I]} \leq C_{E, F(E)}. \quad (2.9)$$

Adding (2.8) and (2.9), the desired bound (2.3) would follow with $F :=$ $F(E) + C_{E, F(E)}$.

Goal A is accomplished by showing that if ϵ is sufficiently small, then $(\tilde{A}, \tilde{\phi})$ is arbitrarily close (i.e., within ϵ^δ) to the frequency truncated solution $(P_{\leq k^*}A, P_{\leq k^*}\phi)$ which has small energy dispersion. For Goal B, the idea is to view (B^{high}, ψ^{high}) as a perturbation around $(\tilde{A}, \tilde{\phi})$. To ensure that $c_0(E)$ is independent of $F(E)$, we rely on two observations: First, by the weak divisibility[10] of the S^1 norm, the interval I can be split into $O_{F(E)}(1)$ many subintervals I_k on each of which we have

$$\|\tilde{A}_0\|_{Y^1[I_k]} + \|\tilde{A}_x\|_{S^1[I_k]} + \|\tilde{\phi}\|_{S^1[I_k]} \lesssim_E 1. \quad (2.10)$$

Second, by conservation of energy for (A, ϕ) and $(\tilde{A}, \tilde{\phi})$, as well as the approximation $(\tilde{A}, \tilde{\phi}) \approx (P_{\leq k^*}A, P_{\leq k^*}\phi)$, it follows that the $\dot{H}_x^1 \times L_x^2$ norm of the data for (B^{high}, ψ^{high}) can be reinitialized to be of size $\lesssim c_0(E)$ on each I_k.

With these two observations in hand, we claim that (B^{high}, ψ^{high}) obeys the following S^1 norm bound on each I_k:

$$\|B_0^{high}\|_{Y^1[I_k]} + \|B_x^{high}\|_{S^1[I_k]} + \|\psi^{high}\|_{S^1[I_k]} \lesssim_E c_0(E) + O_F\left(\epsilon^\delta\right). \quad (2.11)$$

Indeed, in the equation for (B^{high}, ψ^{high}), all nonlinear terms in (B^{high}, ψ^{high}) can be handled by taking $c_0(E) \ll_E 1$ and $\epsilon \ll_F 1$. Furthermore, exploiting small energy dispersion, all linear terms can be made appropriately small except $-2iA_\mu \partial^\mu \psi^{high}$. Nevertheless, the S^1 norm of (A, ϕ) on I can be assumed to be $\lesssim_E 1$ by (2.10) and a bootstrap assumption[11]; hence we can group this term with $\Box$ and use (2.7) (linear well-posedness of $\Box_A \psi^{high}$) to arrive at (2.11). Goal B now follows by summing up this bound on $O_{F(E)}(1)$ intervals.

---

[10] This terminology should be compared with full divisibility, which means that I can be split into a controlled number of subintervals, on each of which the restricted norm is arbitrarily small. Weak divisibility of the S^1 norm is a quick consequence of the energy inequality $\|\psi\|_{S^1[I]} \lesssim \|\psi[0]\|_{\dot{H}_x^1 \times L_x^2} + \|\Box\psi\|_{N[I]}$ and (full) divisibility of the N norm.

[11] More precisely, in proving (2.11) we may assume, using a continuous induction in time, that the same bound holds with a worse constant. Combined with (2.10) this bound is sufficient for ensuring that the S^1 norm of (A, ϕ) is $\lesssim_E 1$.

3 Overview of the proof II: content of the present paper

This section is a continuation of the previous section. Section 3.1 provides an overview of the argument in the present paper, thereby completing the summary of our entire proof of Theorem 1.3. In Sect. 3.2, we give a brief comparison of our approach with that of the recent work [16] of Krieger-Lührmann. Finally, Sect. 3.3 contains an outline of the structure of the remainder of the paper.

3.1 Blow-up analysis

Here we give an overview of the final blow-up analysis of (MKG), which is carried out in the present paper. This part is analogous to [33] for energy critical wave maps. We refer to Sect. 4 for the notation used below.

3.1.1 Main ingredients

In addition to the continuation/scattering criteria established in [26,27] (see Theorems 2.1 and 2.2), our blow-up analysis of (MKG) relies on the following three key ingredients:

– *Monotonicity formula for* (MKG) Besides the conservation of energy, we use the following monotonicity (or Morawetz) formula for (MKG). Let $\rho := \sqrt{t^2 - |x|^2}$ and

$$X_0 := \frac{1}{\rho}(t\partial_t + x \cdot \partial_x)$$

be the normalized scaling vector field. To avoid the degeneracy of ρ on $\partial C = \{t = |x|\}$, we also define the translates

$$\rho_\varepsilon := \sqrt{(t+\varepsilon)^2 - |x|^2}, \quad X_\varepsilon := \frac{1}{\rho_\varepsilon}((t+\varepsilon)\partial_t + x \cdot \partial_x).$$

Given a smooth solution (A, ϕ) to (MKG) on the truncated cone $C_{[\varepsilon,1]}$ satisfying

$$\mathcal{E}_{S_1}[A, \phi] \le E, \quad \mathcal{F}_{\partial C_{[\varepsilon,1]}}[A, \phi] \le \varepsilon^{\frac{1}{2}}E, \quad \mathcal{G}_{S_1}[\phi] \le \varepsilon^{\frac{1}{2}}E,$$

where $\mathcal{F}_{\partial C_{[t_0,t_1]}} := \mathcal{E}_{S_{t_1}} - \mathcal{E}_{S_{t_0}}$ is the energy flux through $\partial C_{[t_0,t_1]}$ and $\mathcal{G}_{S_t} := \frac{1}{t}\int_{S_t}|\phi|^2$, we have

$$\int_{S_1} {}^{(X_\varepsilon)} P_T[A, \phi] \, dx + \iint_{C_{[\varepsilon,1]}} \frac{1}{\rho_\varepsilon} |\iota_{X_\varepsilon} F|^2 + \frac{1}{\rho_\varepsilon} \left| \left(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon} \right) \phi \right|^2 dt \, dx$$

$$\lesssim \int_{S_\varepsilon} {}^{(X_\varepsilon)} P_T[A, \phi] \, dx + E. \tag{3.1}$$

Here ${}^{(X_\varepsilon)} P_T[A, \phi]$ is a non-negative weighted energy density; we refer to Lemma 5.10 for an explicit formula for ${}^{(X_\varepsilon)} P_T[A, \phi]$. We remark that the entire right-hand side of (3.1) is bounded by $\lesssim E$. Finiteness of the space-time integral term 'breaks the scaling' and implies that $\iota_{X_\varepsilon} F$ and $(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon}) \phi$ decay near the tip of the cone C.

– *Strong local compactness result* Given a sequence $(A^{(n)}, \phi^{(n)})$ of solutions whose energy is uniformly small and $\iota_X F^{(n)} \to 0$ and $(\mathbf{D}_X^{(n)} + b) \phi^{(n)} \to 0$ in $L_{t,x}^2$ on a space-time cube for some smooth time-like vector field X and smooth function b, we show that there exists a subsequence which converges strongly in (essentially) $H_{t,x}^1$ in a smaller subcube; see Proposition 6.1 for more details. The proof relies on the initial data excision/gluing technique and the small energy global well-posedness theorem.

– *Triviality of finite energy stationary/self-similar solutions* We say that (A, ϕ) is a stationary solution to (MKG) if for some constant time-like vector field Y

$$\iota_Y F = 0, \quad \mathbf{D}_Y \phi = 0,$$

and that (A, ϕ) is a self-similar solution if

$$\iota_{X_0} F = 0, \quad \left(\mathbf{D}_{X_0} + \frac{1}{\rho} \right) \phi = 0.$$

Using the method of stress tensor, we show that every smooth stationary or self-similar solution with finite energy is trivial (i.e., $F = 0$ and $\phi = 0$); see Propositions 7.1 and 7.2. We also establish a regularity result (Proposition 7.3), which says that all stationary and self-similar solutions arising from the above strong local compactness result (Proposition 6.1) are smooth.

With these in mind, we now sketch the blow-up analysis of (MKG), which is performed in full detail in Sect. 8.

3.1.2 Finite time blow-up/non-scattering scenarios and initial reduction

Suppose that the conclusion of Theorem 1.3 fails for a smooth finite energy data (a, e, f, g) in the forward time direction. Then the corresponding smooth

solution either blows up in finite time, or does not scatter as $t \to \infty$. The first step of the blow-up analysis is to construct in both scenarios a sequence of global Coulomb solutions $(A^{(n)}, \phi^{(n)})$ on $[\varepsilon_n, 1] \times \mathbb{R}^4$ (where $\varepsilon_n \to 0$) obeying the following properties:

– *Bounded energy in the cone* $\mathcal{E}_{S_t}[A^{(n)}, \phi^{(n)}] \leq E$ for every $t \in [\varepsilon_n, 1]$
– *Small energy outside the cone* $\mathcal{E}_{(\{t\} \times \mathbb{R}^4) \setminus S_t}[A^{(n)}, \phi^{(n)}] \ll E$ for every $t \in [\varepsilon_n, 1]$
– *Decaying flux on* ∂C $\mathcal{F}_{[\varepsilon_n, 1]}[A^{(n)}, \phi^{(n)}] + \mathcal{G}_{S_1}[\phi^{(n)}] \leq \varepsilon_n^{\frac{1}{2}} E$,
– *Pointwise concentration at* $t = 1$ There exist $k_n \in \mathbb{Z}$ and $x_n \in \mathbb{R}^4$ such that

$$2^{-k_n} |\zeta_{2^{-k_n}} * \phi^{(n)}(1, x_n)| + 2^{-2k_n} |\zeta_{2^{-k_n}} * \mathbf{D}_t^{(n)} \phi^{(n)}(1, x_n)| > \mathbf{e} \quad (3.2)$$

for some $\mathbf{e} = \mathbf{e}(E) > 0$.

Here ζ is a smooth function supported in the unit ball $B_1(0)$ and $\zeta_{2^{-k}}(x) := 2^{4k} \zeta(2^k x)$. In view of the next step, we require ζ to be non-negative. See Lemma 8.4 for details.

Key to this construction are Theorems 2.1 and 2.2, which provide detailed information about finite time blow-up or non-scattering scenarios. In particular, the tip of the cone C is the point of energy concentration (which exists by Theorem 2.1) in the finite time blow-up case. Pointwise concentration at $t = 1$ follows from the failure of the energy dispersion bound in Theorem 2.2. Decaying flux on ∂C is a consequence of the local conservation of energy and localized Hardy's inequality; see Lemma 5.2 and Corollary 5.3. Smallness of the energy outside the cone is achieved using the initial data excision/gluing technique in the finite time blow-up case; in the non-scattering case, this property is trivial to establish.

3.1.3 Elimination of the null concentration scenario

Thanks to the above properties, we may apply the monotonicity formula (3.1) to each solution in the sequence $(A^{(n)}, \phi^{(n)})$. Using the weighted energy term (i.e., the first term on the left-hand side) in (3.1), we show in Lemma 8.7 that the null concentration scenario (i.e., $|x_n| \to 1$ and $k_n \to \infty$) is impossible. Unlike in the case of wave maps [33], however, the weighted energy involves the covariant derivatives $\mathbf{D}_\mu^{(n)} \phi^{(n)} = \partial_\mu \phi^{(n)} + i A_\mu^{(n)} \phi^{(n)}$, and the term involving $A^{(n)}$ could be problematic. We avoid this issue by first working with the gauge invariant amplitude $|\phi^{(n)}|$, for which we have the diamagnetic inequality

$$|X^\mu \partial_\mu |\phi^{(n)}|| \leq |\mathbf{D}_X \phi^{(n)}|$$

in the sense of distributions, for any smooth vector field X. We then transfer the bound to $\phi^{(n)}$ using the inequality

$$|\zeta_{2^{-k}} * \phi^{(n)}| \leq \zeta_{2^{-k}} * |\phi^{(n)}|,$$

which holds if ζ is chosen to be non-negative.

3.1.4 Nontrivial energy in a time-like region

The absence of the null concentration scenario implies the following uniform lower bound for $\phi^{(n)}$ away from the boundary at $t = 1$: There exist $E_1 = E_1(E) > 0$ and $\gamma_1 = \gamma_1(E) \in (0, 1)$ such that

$$\int_{S_1^{1-\gamma_1}} \sum_{\mu=0}^{4} |\mathbf{D}_\mu^{(n)} \phi^{(n)}|^2 + \frac{1}{r^2} |\phi^{(n)}|^2 \, dx \geq E_1. \tag{3.3}$$

See Lemma 8.9. Using a localized version of the monotonicity formula (3.1), this lower bound can be propagated towards $t = 0$. More precisely, there exist $E_2 = E_2(E)$ and $\gamma_2 = \gamma_2(E) \in (0, 1)$ and $E_2 = E_2(E) > 0$ such that

$$\int_{S_t^{(1-\gamma_2)t}} {}^{(X_0)} P_T \left[A^{(n)}, \phi^{(n)} \right] dx \geq E_2 \quad \text{for all } t \in \left[\varepsilon_n^{\frac{1}{2}}, \varepsilon_n^{\frac{1}{4}} \right]. \tag{3.4}$$

3.1.5 Final rescaling

Thanks to the space-time integral term in (3.1), $(A^{(n)}, \phi^{(n)})$ obeys

$$\iint_{C_{[\varepsilon_n, 1]}} \frac{1}{\rho_{\varepsilon_n}} |\iota_{X_{\varepsilon_n}} F^{(n)}|^2 + \frac{1}{\rho_{\varepsilon_n}} \left| \left(\mathbf{D}_{X_{\varepsilon_n}}^{(n)} + \frac{1}{\rho_{\varepsilon_n}} \right) \phi^{(n)} \right|^2 dt dx \lesssim E.$$

which implies an integrated decay of $\iota_{X_{\varepsilon_n}} F^{(n)}$ and $(\mathbf{D}_{X_{\varepsilon_n}}^{(n)} + \frac{1}{\rho_{\varepsilon_n}}) \phi^{(n)}$ near the tip of the cone C. Applying the pigeonhole principle and rescaling, we obtain a new sequence of solutions which is asymptotically self-similar. More precisely, there exist a sequence of solutions on $[1, T_n] \times \mathbb{R}^4$ (where $T_n \to \infty$) to (MKG), which we still denote by $(A^{(n)}, \phi^{(n)})$, obeying the following properties (see Lemma 8.11):

– *Bounded energy in the cone* $\mathcal{E}_{S_t}[A^{(n)}, \phi^{(n)}] \leq E$ *for every* $t \in [1, T_n]$,
– *Small energy outside the cone* $\mathcal{E}_{\{t\} \times \mathbb{R}^4 \setminus S_t}[A^{(n)}, \phi^{(n)}] \ll E$ *for every* $t \in [1, T_n]$,

– *Nontrivial energy in a time-like region* For every $t \in [1, T_n]$ we have

$$\int_{S_t^{(1-\gamma_2)t}} {}^{(X_0)} P_T \left[A^{(n)}, \phi^{(n)} \right] dx \geq E_2, \tag{3.5}$$

– *Asymptotic self-similarity* For every compact subset K of the interior of $C_{[1,\infty)}$, we have

$$\iint_K |\iota_{X_0} F^{(n)}|^2 + \left| \left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 dt dx \to 0 \quad \text{as } n \to \infty. \tag{3.6}$$

3.1.6 Extraction of concentration scales and compactness/rigidity argument

Let $(A^{(n)}, \phi^{(n)})$ be a sequence obtained by the final rescaling argument. Using a combinatorial argument, we show in Lemma 8.12 that one of the following two scenarios holds:

A. Either we can identify a sequence of points and decreasing scales at which energy concentrates, or
B. There is a uniform non-concentration of energy.

In Scenario A we obtain a fixed number $r > 0$ and a sequence of times $t_n \to t_0$, points $x_n \to x_0$ and scales $r_n \to 0$ such that

$$\sup_{x \in B_r(x_n)} \mathcal{E}_{\{t_n\} \times B_{r_n}(x)} \left[A^{(n)}, \phi^{(n)} \right]$$

is uniformly small but nontrivial, and

$$\frac{1}{4r_n} \int_{t_n-2r_n}^{t_n+2r_n} \int_{B_r(x_n)} |\iota_Y F^{(n)}|^2 + |\mathbf{D}_Y^{(n)} \phi^{(n)}|^2 dt dx \to 0 \quad \text{as } n \to \infty.$$

where $Y = X_0(t_0, x_0)$. Applying Proposition 6.1, we obtain as a limit a nontrivial finite energy solution to (MKG) which is stationary with respect to Y. As discussed above, however, such solutions do not exist.

In Scenario B we can cover each truncated cone $\widetilde{C}_j := C_{[1/2,\infty)}^{1/2} \cap \{2^j \leq t < 2^{j+1}\}$ with spatial balls of radius $r = r(j)$, on each of which the energy of $(A^{(n)}, \phi^{(n)})$ is uniformly small and

$$\iint_{\widetilde{C}_j} |\iota_{X_0} F^{(n)}|^2 + \left| \left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 dt dx \to 0 \quad \text{as } n \to \infty.$$

Hence we are again in position to apply Proposition 6.1 and extract a finite energy self-similar solution to (MKG) on $C_{[1/2,\infty)}^{1/2}$. By self-similarity, this

limit easily extends to the whole forward cone C. By (3.5) this limit is necessarily nontrivial, which contradicts the triviality of finite energy self-similar solutions.

In conclusion, we have seen that neither of the two scenarios can hold, which is a contradiction. This completes the proof of the main theorem.

3.2 Comparison with the approach of [16]

The principal difference between the present work and [16] is that the latter follows the Kenig-Merle concentration compactness/rigidity scheme [9] for establishing global well-posedness and scattering. Roughly, this scheme consists of two steps: First, assuming that the conclusion fails, one constructs a blow-up solution with the minimal energy. Second, one derives a contradiction by playing various conservation laws and monotonicity formulae of (MKG) against special compactness properties of the minimal blow-up solution. As an immediate corollary, this approaches yields some additional information about the solutions, such as an a-priori bound on the S^1 norm in terms of the energy. On the other hand, as we explain below, the execution of this scheme in the presence of a non-perturbative paradifferential nonlinearity faces considerable difficulty, a large part of which is avoided in the present work.

The main ingredient for construction of a minimal blow-up solution is the concentration compactness phenomenon, or nonlinear profile decomposition, for (MKG) in the global Coulomb gauge, whose proof takes up the majority of [16]. This concept was first introduced in the context of elliptic PDEs by Lions [21] and was adapted to the semilinear wave equation $\Box u = |u|^4 u$ on $\mathbb{R}^{1+3}$ by Bahouri-Gérard [1]. It roughly states that any sequence of solutions with uniformly bounded energy can be decomposed (after passing to a subsequence, on a suitable time interval) into the superposition of solutions modulated by the non-compact symmetries of the problem (called profiles) and an error which can be made arbitrarily small in an appropriate norm weaker than energy.

Key to the proof of concentration compactness in [1] is the asymptotic vanishing of the interaction among different profiles, whose frequency and space-time supports are diverging from each other. However, such a statement partly fails for (MKG), due to the non-perturbative effect of the low frequency part of the solution on the high frequency part through the paradifferential nonlinearity. In [16], this difficulty is overcome by performing an induction on frequency, where one carefully builds a profile decomposition in the order of increasing frequency. This delicate procedure, which originated from the earlier work of Krieger-Schlag [17], necessitates several ideas not used in our approach, such as a process for extracting linear profiles using the paradifferential magnetic wave equation and a uniform dispersive estimate for such a equation (see [16, Sect. 7.5]).

Another notable difference between this paper and [16] is the conservation laws and monotonicity formulae used in the proof. While the present paper relies only on the energy conservation and the Morawetz-type monotonicity formula (3.1), [16] uses in addition momenta conservation and a virial-type monotonicity formula for (MKG), which are of independent interest.

3.3 Structure of the present paper

The remainder of the paper is structured as follows.

Section 4

We provide the setup for our arguments to follow. In particular, we precisely state the results that we need from the other papers of the series [26,27] in Sect. 4.5.

Section 5

We state and prove all the conservation laws and monotonicity formulae that are used in this paper.

Section 6

We use the small energy global well-posedness theorem (Theorem 4.1) and the technique of initial data excision/gluing to prove a strong local compactness statement (Proposition 6.1) that we rely on in our blow-up analysis. We also formulate a notion of weak solutions to (MKG) and their local descriptions (weak compatible pairs), which naturally arise as limits from Proposition 6.1.

Section 7

We show that there does not exist any nontrivial stationary or self-similar solutions to (MKG) with finite energy. We also prove regularity theorems for weak stationary or self-similar solutions to (MKG) considered in Sect. 6.

Section 8

We finally carry out the blow-up analysis as outlined in Sect. 3.1, thereby completing the proof of global well-posedness and scattering of (MKG).

4 Preliminaries

4.1 Notation for constants and asymptotics

Throughout the paper we use C for a general positive constant, which may vary from line to line. For a constant C that depends on, say, E, we write $C = C(E)$. We write $A \lesssim B$ when there exists a constant $C > 0$ such that $A \leq CB$. When the implicit constant should be regarded as small, we write $A \ll B$. The dependence of the constant is specified by a subscript, e.g., $A \lesssim_E B$. We write $A \approx B$ when both $A \lesssim B$ and $B \lesssim A$ hold.

4.2 Coordinate systems on $\mathbb{R}^{1+4}$

Several different coordinate systems on $\mathbb{R}^{1+4}$ will be used in this paper. A basic choice, which has already been mentioned in the introduction, is the rectilinear coordinates $(x^0, x^1, \ldots, x^4)$ on $\mathbb{R}^{1+4}$, in which the Minkowski metric takes the diagonal form $\mathbf{m} = -(dx^0)^2 + (dx^1)^2 + \cdots + (dx^4)^2$. Alternatively, we will often write $t = x^0$ and $x = (x^1, \ldots, x^4)$ as well. We reserve the greek indices $\mu, \nu, \ldots$ for expressions in the rectilinear coordinates, and the latin indices $j, k, \ell, \ldots$ expressions only in terms of the spatial coordinates x^1, x^2, x^3, x^4.

We also introduce the polar coordinates (t, r, Θ) on $\mathbb{R}^{1+4}$, where

$$r = |x|, \quad \Theta = \frac{x}{|x|} \in \mathbb{S}^3,$$

and the null coordinates (u, v, Θ), defined by

$$u = t - r, \quad v = t + r.$$

We can furthermore specify a spherical coordinate system for Θ, but it will not be necessary. We also define the null vector fields $L, \underline{L}$ as

$$L = \partial_t + \partial_r = 2\partial_v, \quad \underline{L} = \partial_t - \partial_r = 2\partial_u.$$

In these coordinates, the metric takes the form

$$\mathbf{m} = -dt^2 + dr^2 + r^2 g_{\mathbb{S}^3} = -du dv + r^2(u, v) g_{\mathbb{S}^3}.$$

where $g_{\mathbb{S}^3}$ is the standard metric on $\mathbb{S}^3$ in the coordinates Θ.

Finally, we will also use the hyperbolic polar coordinates (in short, hyperbolic coordinates) (ρ, y, Θ) on the future light cone $C_{(0,\infty)} = \{(t, r, \Theta) : 0 \leq r < t\}$ (see below), where

$$\rho = \sqrt{t^2 - r^2}, \quad y = \tanh^{-1}(r/t).$$

The Minkowski metric takes the form

$$\mathbf{m} = -\mathrm{d}\rho^2 + \rho^2 \left(\mathrm{d}y^2 + \sinh^2 y \, g_{\mathbb{S}^3} \right).$$

Every constant ρ hypersurface $\mathcal{H}_\rho$ is isometric to the simply connected space of constant sectional curvature $-\frac{1}{\rho^2}$; in particular, $\mathcal{H}_1$ is the hyperboloidal model for the hyperbolic 4-space $\mathbb{H}^4$. Using the coordinates (y, Θ), the metric on $\mathbb{H}^4$ can be written as

$$g_{\mathbb{H}^4} = \mathrm{d}y^2 + \sinh^2 y \, g_{\mathbb{S}^3}.$$

4.3 Geometric notation

To ease the transition from one coordinate system to another, we shall use the tensor formalism. We will denote by ∇ the Levi-Civita connection on $\mathbb{R}^{1+4}$ to distinguish from coordinate vector fields ∂_μ. The gauge covariant connection associated to A for $\mathbb{C}$-valued tensors takes the form $\mathbf{D} = \nabla + iA$. Similarly, we shall denote the Levi-Civita connection on $\mathbb{H}^4$ by $\nabla_{\mathbb{H}^4}$, and the gauge covariant connection by $\mathbf{D}_{\mathbb{H}^4} = \nabla_{\mathbb{H}^4} + iA$. We use the bold latin indices $\mathbf{a}, \mathbf{b}, \ldots$ for expressions in a general coordinate system. We also employ the usual convention of raising and lowering indices using the Minkowski metric $\mathbf{m}$, and summing up repeated upper and lower indices.

We now introduce some notation for geometric subsets of $\mathbb{R}^{1+4}$ and $\mathbb{R}^4$. The forward light cone

$$C := \{(t, x) : 0 < t < \infty, |x| \le t\}$$

will play a central role in this paper. For $t_0 \in \mathbb{R}$ and $I \subset \mathbb{R}$, we define

$$
\begin{aligned}
C_I &:= \{(t, x) : t \in I, |x| \le t\}, & \partial C_I &:= \{(t, x) : t \in I, |x| = t\}, \\
S_{t_0} &:= \{(t, x) : t = t_0, |x| \le t\}, & \partial S_{t_0} &:= \{(t, x) : t = t_0, |x| = t\}.
\end{aligned}
$$

For $\delta \in \mathbb{R}$, we define the translated cones

$$C^\delta := \{(t, x) : \max\{0, \delta\} \le t < \infty, |x| \le t - \delta\}.$$

The corresponding objects $C_I^\delta, \partial C_I^\delta, S_{t_0}^\delta$ and $\partial S_{t_0}^\delta$ are defined in the obvious manner.

We also define $B_r(x)$ to be the ball of radius r centered at x in $\mathbb{R}^4$.

4.4 Frequency projections and function spaces

Let $m_{\leq 0}$ be a smooth cutoff that equals 1 on $\{r \leq 1\}$ and 0 on $\{r \geq 2\}$. For $k \in \mathbb{Z}$, we define

$$m_{\leq k}(r) := m_{\leq 0}(r/2^k), \quad m_k(r) := m_{\leq k}(r) - m_{\leq k-1}(r).$$

so that $\operatorname{supp} m_k \subseteq \{2^{k-1} \leq r \leq 2^{k+1}\}$ and $\sum_k m_k(r) = 1$. We introduce the Littlewood-Paley projections P_k, Q_j and S_ℓ, which are used in this paper:

$$P_k \varphi = \mathcal{F}^{-1}[m_k(|\xi|)\mathcal{F}[\varphi]],$$
$$Q_j \varphi = \mathcal{F}^{-1}[m_j(||\tau| - |\xi||)\mathcal{F}[\varphi]],$$
$$S_\ell \varphi = \mathcal{F}^{-1}[m_\ell(|(\tau, \xi)|)\mathcal{F}[\varphi]],$$

where $\mathcal{F}$ [resp. $\mathcal{F}^{-1}$] is the [resp. the inverse] space-time Fourier transform.

Given a normed space X of function on $\mathbb{R}^{1+4}$, we define the restriction space $X(\mathcal{O})$ on a measurable subset $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ by the norm

$$\|\varphi\|_{X(\mathcal{O})} := \inf_{\psi = \varphi \text{ on } \mathcal{O}} \|\psi\|_{X(\mathbb{R}^{1+4})}.$$

In application, the set $\mathcal{O}$ is often an open set with (piecewise) smooth boundary, and hence there exists a bounded linear extension operator from $X(\mathcal{O})$ to $X(\mathbb{R}^{1+4})$ for many standard function spaces X (e.g., $X = H^1$).

4.5 Results from previous papers

Here we give precise statements of results from [19] and the first two papers in the sequence [26,27], which are used in the present paper. Given a measurable subset $S \subseteq \{t\} \times \mathbb{R}^4$ for some t, we define the energy of a pair (A, ϕ) on S by

$$\mathcal{E}_S[A, \phi] := \int_S \frac{1}{2} \sum_{0 \leq \mu < \nu \leq 4} |F_{\mu\nu}|^2 + \frac{1}{2} \sum_{\mu=0}^{4} |\mathbf{D}_\mu \phi|^2 \, dx.$$

Accordingly, for a measurable subset $S \subseteq \mathbb{R}^4$, we define

$$\mathcal{E}_S[a, e, f, g] := \int_S \frac{1}{2} \sum_{1 \leq j < k \leq 4} |(da)_{jk}|^2 + \frac{1}{2} \sum_{j=1}^{4} |e_j|^2$$
$$+ \frac{1}{2} \sum_{j=1}^{4} |\mathbf{D}_j f|^2 + \frac{1}{2} |g|^2 \, dx.$$

The following is the main theorem of [19].

Theorem 4.1 (Small energy global well-posedness in global Coulomb gauge)
There exists $\epsilon_* > 0$ such that the following holds. Let (a, e, f, g) be an $\mathcal{H}^1$
initial data set on $\mathbb{R}^4$ satisfying the global Coulomb gauge condition $\partial^\ell a_\ell = 0$,
whose energy does not exceed ϵ_*^2, i.e.,

$$\mathcal{E}_{\mathbb{R}^4}[a, e, f, g] \leq \epsilon_*^2. \tag{4.1}$$

(1) *Then there exists a unique $C_t \mathcal{H}^1$ admissible solution (A, ϕ) to (MKG)*
   on $\mathbb{R}^{1+4}$ satisfying the global Coulomb gauge condition $\partial^\ell A_\ell = 0$ with
   (a, e, f, g) as its initial data at $t = 0$, i.e., $(A_j, F_{0j}, \phi, \mathbf{D}_t \phi) \restriction_{\{t=0\}} =$
   (a_j, e_j, f, g).
(2) *Moreover, (A, ϕ) obeys the S^1 norm bound*

$$\|A_0\|_{Y^1(\mathbb{R}^{1+4})} + \|A_x\|_{S^1(\mathbb{R}^{1+4})} + \|\phi\|_{S^1(\mathbb{R}^{1+4})} \lesssim \|(a, e, f, g)\|_{\mathcal{H}^1}. \tag{4.2}$$

(3) *If the initial data set (a, e, f, g) is more regular, then so is the solution*
   (A, ϕ); in particular, if (a, e, f, g) is classical, then (A, ϕ) is a classical
   solution to (MKG).
(4) *Finally, given a sequence $(a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)}) \in \mathcal{H}^1(\mathbb{R}^4)$ of Coulomb*
   initial data sets such that $\mathcal{E}[a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)}] \leq \epsilon_*^2$ and $(a^{(n)}, e^{(n)},$
   $f^{(n)}, g^{(n)}) \to (a, e, f, g)$ in $\mathcal{H}^1(\mathbb{R}^4)$, we have

$$\|A_0^{(n)} - A_0\|_{Y^1(I \times \mathbb{R}^4)} + \|A_x^{(n)} - A_x\|_{S^1(I \times \mathbb{R}^4)} + \|\phi^{(n)} - \phi\|_{S^1(I \times \mathbb{R}^4)} \to 0 \tag{4.3}$$

   as $n \to \infty$, for every compact interval $I \subseteq \mathbb{R}$.

The first statement can be found directly in the main theorem of [19]. For the
proof of the remaining statements, see [19, Sect. 5].

Remark 4.2 For the purpose of the present paper, the precise structure of the
norm S^1 is not necessary. Instead, we rely on the following embedding prop-
erty:

$$\|\partial_{t,x}\phi\|_{L_t^\infty L_x^2} + \|\Box\phi\|_{L_t^2 \dot{H}_x^{-\frac{1}{2}}} \lesssim \|\phi\|_{S^1},$$

where all norms are taken on $\mathbb{R}^{1+4}$; see [19, Sect. 3] or [27, Sect. 3]. On the
other hand, the definition of the Y^1 norm is rather simple:

$$\|A\|_{Y^1}^2 := \|\partial_{t,x} A\|_{L_t^\infty L_x^2}^2 + \|\partial_{t,x} A\|_{L^2 \dot{H}_x^{\frac{1}{2}}}^2$$

where all norms are again taken on $\mathbb{R}^{1+4}$. Furthermore, S^1 and Y^1 are closed under multiplication by $\eta \in C_0^\infty(\mathbb{R}^{1+4})$, i.e., $\eta S^1(\mathbb{R}^{1+4}) \subseteq S^1(\mathbb{R}^{1+4})$ and $\eta Y^1(\mathbb{R}^{1+4}) \subseteq Y^1(\mathbb{R}^{1+4})$; we refer to [26, Sects. 6 and 7].

Given a positive number $E \gtrsim \epsilon_*$ and an $\mathcal{H}^1$ initial data set (a, e, f, g) on $\mathbb{R}^4$ with energy $\mathcal{E}[a, e, f, g] \leq E$, we define its energy concentration scale $r_c = r_c[a, e, f, g]$ (with respect to energy E), in terms of the function $\delta_0(E, \epsilon_*^2) = c\epsilon_*^2 \min\{1, \epsilon_*^4 E^{-2}\}$ with a small universal constant c, by

$$r_c = r_c(E)[a, e, f, g] := \sup\{r \geq 0 : \forall x \in \mathbb{R}^4, \ \mathcal{E}_{B_r(x)}[a, e, f, g]$$
$$< \delta_0\left(E, \epsilon_*^2\right)\}. \tag{4.4}$$

The following is the main result of [26].

Theorem 4.3 (Large energy local well-posedness theorem in global Coulomb gauge) *Let (a, e, f, g) be an $\mathcal{H}^1$ initial data set satisfying the global Coulomb gauge condition $\partial^\ell a_\ell = 0$ with energy $\mathcal{E}[a, e, f, g] \leq E$. Let $r_c = r_c[a, e, f, g]$ be defined as above. Then the following statements hold:*

(1) Existence and uniqueness. *There exists a unique admissible $C_t \mathcal{H}^1$ solution (A, ϕ) to (MKG) on $[-r_c, r_c] \times \mathbb{R}^4$ satisfying the global Coulomb gauge condition with (a, e, f, g) as its initial data.*
(2) A-priori S^1 regularity. *We have the additional regularity properties*

$$A_0 \in Y^1[-r_c, r_c], \quad A_x, \phi \in S^1[-r_c, r_c].$$

(3) Persistence of regularity. *If the initial data set (a, e, f, g) is more regular, then so is the solution (A, ϕ); in particular, the solution (A, ϕ) is classical if (a, e, f, g) is classical.*
(4) Continuous dependence. *Consider a sequence $(a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)})$ of $\mathcal{H}^1$ Coulomb initial data sets such that $(a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)}) \to (a, e, f, g)$ in $\mathcal{H}^1$. Then the lifespan of $(A^{(n)}, \phi^{(n)})$ eventually contains $[-r_c, r_c]$, and we have*

$$\|A_0 - A_0^{(n)}\|_{Y^1[-r_c, r_c]} + \|\left(A_x - A_x^{(n)}, \phi - \phi^{(n)}\right)\|_{S^1[-r_c, r_c]} \to 0$$
as $n \to \infty$.

We also state the initial data excision/gluing result from [26], which is used in several places in the present paper. Given a measurable subset $O \subseteq \mathbb{R}^4$, the $\mathcal{H}^1(O)$ norm is defined as the restriction of the $\mathcal{H}^1(\mathbb{R}^4)$ norm to O, and the space $\mathcal{H}^1(O)$ consists of all initial data sets on O with finite $\mathcal{H}^1(O)$ norm.

Proposition 4.4 (Excision and gluing of initial data sets) *Let $B = B_{r_0}(x_0) \subseteq \mathbb{R}^4$. Then there exists an operator E^{ext} from $\mathcal{H}^1(2B \backslash \overline{B})$ to $\mathcal{H}^1(\mathbb{R}^4 \backslash \overline{B})$ satisfying the following properties.*

(1) *Extension property:*

$$E^{\text{ext}}[a, e, f, g] = (a, e, f, g) \quad \text{on the annulus } \frac{3}{2}B \backslash \overline{B}.$$

(2) *Uniform bounds:*

$$\|E^{\text{ext}}[a, e, f, g]\|_{\mathcal{H}^1(\mathbb{R}^4 \backslash \overline{B})} \lesssim \|(a, e, f, g)\|_{\mathcal{H}^1(2B \backslash \overline{B})} \tag{4.5}$$

$$\mathcal{E}_{\mathbb{R}^4 \backslash \overline{B}}[E^{\text{ext}}[a, e, f, g]] \lesssim \|\frac{1}{|x - x_0|} f\|_{L_x^2(2B \backslash \overline{B})}^2 + \mathcal{E}_{2B \backslash \overline{B}}[a, e, f, g]. \tag{4.6}$$

(3) *Regularity: The operator E^{ext} is continuous from $\mathcal{H}^1(2B \backslash \overline{B})$ to $\mathcal{H}^1(\mathbb{R}^4 \backslash \overline{B})$. Moreover, if (a, e, f, g) is classical, then so is $E^{\text{ext}}[a, e, f, g]$.*

In order to gain control of the first norm on the right in (4.6), we will repeatedly use the following improvement of the classical Hardy inequality, which is a consequence of a result proved in [26], Lemma 6.5:

Lemma 4.5 *Let $\sigma \geq 2$. Then for any ball B of radius r in $\mathbb{R}^4$ we have the bounds*

$$r^{-1}\|f\|_{L_x^2(2B)} \lesssim \|\mathbf{D}_x f\|_{L_x^2(\sigma B)} + \sigma^{-1}\|\mathbf{D}_x f\|_{L_x^2(\mathbb{R}^4 \backslash \overline{\sigma B})} \tag{4.7}$$

$$r^{-1}\|f\|_{L_x^2(2B \backslash \overline{B})} \lesssim \|\mathbf{D}_x f\|_{L_x^2(\sigma B \backslash \overline{B})} + \sigma^{-1}\|\mathbf{D}_x f\|_{L_x^2(\mathbb{R}^4 \backslash \overline{\sigma B})} \tag{4.8}$$

Furthermore, we state the local geometric uniqueness result from [26], which we use in this paper to construct compatible pairs. For a ball $B = \{t_0\} \times B_{r_0}(x_0) \subseteq \{t_0\} \times \mathbb{R}^4$, we define its future domain of dependence $\mathcal{D}^+(B)$ to be the set

$$\mathcal{D}^+(B) := \left\{(t, x) \in \mathbb{R}^{1+4} : t_0 \leq t < r_0, \ |x - x_0| < t - t_0\right\}.$$

Given a measurable subset $O \subseteq \mathbb{R}^4$, the space $\mathcal{G}^2(O)$ consists of locally integrable gauge transformations such that the following semi-norm is finite:

$$\|\chi\|_{\mathcal{G}^2(O)} := \|\partial_x \chi\|_{L_x^4(O)} + \|\partial_x^{(2)} \chi\|_{L_x^2(O)}.$$

Given a measurable subset $\mathcal{O} \subseteq \mathbb{R}^{1+4}$, define $\mathcal{O}_t := \mathcal{O} \cap (\{t\} \times \mathbb{R}^4)$ and $I(\mathcal{O}) := \{t \in \mathbb{R} : \mathcal{O}_t \neq \emptyset\}$. Note that $I(\mathcal{O})$ is measurable and $\mathcal{O}_t$ is measurable for almost every t. Accordingly, we define the space $C_t \mathcal{G}^2(\mathcal{O})$ by the semi-norm

$$\|\chi\|_{C_t \mathcal{G}^2(\mathcal{O})} := \operatorname*{ess\,sup}_{t \in I(\mathcal{O})} \left(\|\chi\|_{\dot{H}_x^2 \cap \dot{W}_x^{1,4} \cap \mathrm{BMO}(\mathcal{O}_t)} \right.$$

$$\left. + \|\partial_t \chi\|_{\dot{H}_x^1 \cap L_x^4(\mathcal{O}_t)} + \|\partial_t^2 \chi\|_{L_x^2(\mathcal{O}_t)} \right).$$

Proposition 4.6 (Local geometric uniqueness among admissible solutions) *Let $T_0 > 0$ and let $B \subset \mathbb{R}^4$ be an open ball. Consider $C_t \mathcal{H}^1$ admissible solutions (A, ϕ), (A', ϕ') on the region*

$$\mathcal{D} := \mathcal{D}^+(\{0\} \times B) \cap ([0, T_0) \times \mathbb{R}^4).$$

Suppose that the respective initial data (a, e, f, g) and (a', e', f', g') are gauge equivalent on B, i.e., there exists $\underline{\chi} \in \mathcal{G}^2(B)$ such that $(a, e, f, g) = (a' - \mathrm{d}\underline{\chi}, e', e^{i\underline{\chi}} f', e^{i\underline{\chi}} g')$. Then there exists a unique gauge transformation $C_t \mathcal{G}^2(\mathcal{D})$ such that $\chi \restriction_{\{0\} \times B} = \underline{\chi}$ and

$$(A, \phi) = (A' - \mathrm{d}\chi, e^{i\chi} \phi') \quad on \ \mathcal{D}.$$

We now pass to results from [27]. Given an interval $I \subseteq \mathbb{R}$, we define the *energy dispersion* of a function ϕ on $I \times \mathbb{R}^4$ by

$$ED[\phi](I) := \sup_{k \in \mathbb{Z}} \left(2^{-k} \|P_k \phi\|_{L_{t,x}^\infty(I \times \mathbb{R}^4)} + 2^{-2k} \|P_k(\partial_t \phi)\|_{L_{t,x}^\infty(I \times \mathbb{R}^4)} \right) \quad (4.9)$$

The main theorem of [27] is as follows.

Theorem 4.7 (Energy dispersed regularity theorem) *For each $E > 0$ there exist positive numbers $\epsilon = \epsilon(E)$ and $F = F(E)$ such that the following holds. Let $I \subseteq \mathbb{R}$ be an open interval, and let (A, ϕ) be an admissible $C_t \mathcal{H}^1$ solution to* (MKG) *on $I \times \mathbb{R}^4$ in the global Coulomb gauge $\partial^\ell A_\ell = 0$ with energy not exceeding E, i.e.,*

$$\mathcal{E}_{\{t\} \times \mathbb{R}^4}[A, \phi] \le E \quad for \ every \ t \in I. \quad (4.10)$$

If, furthermore, the energy dispersion of ϕ on $I \times \mathbb{R}^4$ is less than or equal to $\epsilon(E)$, i.e.,

$$ED[\phi](I) \le \epsilon(E), \quad (4.11)$$

then the following a-priori estimate for (A, ϕ) on $I \times \mathbb{R}^4$ holds:

$$\|A_0\|_{Y^1[I]} + \|A_x\|_{S^1[I]} + \|\phi\|_{S^1[I]} \le F(E). \quad (4.12)$$

We also state an continuation and scattering result for Coulomb solutions with finite S^1 norm, which is proved in [27].

Theorem 4.8 (Continuation and scattering of solutions with finite S^1 norm) *Let $0 < T_+ \leq \infty$ and (A, ϕ) an admissible $C_t \mathcal{H}^1$ solution to (MKG) on $[0, T_+) \times \mathbb{R}^4$ in the global Coulomb gauge which obeys the bound*

$$\|A_0\|_{Y^1([0,T_+)\times\mathbb{R}^4)} + \sup_{j=1,\ldots,4} \|A_j\|_{S^1([0,T_+)\times\mathbb{R}^4)} + \|\phi\|_{S^1([0,T_+)\times\mathbb{R}^4)} < \infty.$$

Then the following statements hold.

(1) *If $T_+ < \infty$, then (A, ϕ) extends to an admissible $C_t \mathcal{H}^1$ solution with finite S^1 norm past T_+.*
(2) *If $T_+ = \infty$, then (A_x, ϕ) scatters as $t \to \infty$ in the following sense: There exist a solution $(A_x^{(\infty)}, \phi^{(\infty)})$ to the system*

$$\begin{cases} \Box A_j^{(\infty)} = 0, \\ \left(\Box + 2i A_\ell^{free} \partial^\ell \right) \phi^{(\infty)} = 0, \end{cases}$$

with initial data $A_x^{(\infty)}[0], \phi^{(\infty)}[0] \in \dot{H}_x^1 \times L_x^2$ such that

$$\sup_{j=1,\ldots,4} \|A_j[t] - A_j^{(\infty)}[t]\|_{\dot{H}_x^1 \times L_x^2}$$
$$+ \|\phi[t] - \phi^{(\infty)}[t]\|_{\dot{H}_x^1 \times L_x^2} \to 0 \quad as\ T \to \infty.$$

The above statement holds with A_x^{free} the solution to the homogeneous wave equation with any of the data $A_x[0]$ or $A_x^{(\infty)}[0]$.

Analogous statements hold in the past time direction as well.

5 Conservation laws and monotonicity formulae

In this section, we derive key conservation laws and monotonicity formulae that will serve as a basis for proving regularity and scattering. We begin by describing the main results, deferring their proofs until later in the section. We emphasize that all statements in this section apply to admissible $C_t \mathcal{H}^1$ solutions to (MKG), unless otherwise stated.

One of the fundamental conservation laws for (MKG) is that of the standard energy: Given an admissible $C_t \mathcal{H}^1$ solution (A, ϕ) to (MKG) on $I \times \mathbb{R}^4$, for $t_0, t_1 \in I$ we have

$$\mathcal{E}_{\{t_0\}\times\mathbb{R}^4}[A, \phi] = \mathcal{E}_{\{t_1\}\times\mathbb{R}^4}[A, \phi]. \tag{5.1}$$

For self-similar solutions, finite energy condition translates to a weighted L^2 estimate on $\mathcal{H}_\rho$. This estimate will be used to show that they must in fact be trivial.

Proposition 5.1 *Let (A, ϕ) be a smooth solution to* (MKG) *on $C_{(0,\infty)}$ with finite energy, i.e., there exists $E > 0$ such that*

$$\sup_{t \in (0,\infty)} \mathcal{E}_{S_t}[A, \phi] \leq E < \infty.$$

Suppose furthermore that (A, ϕ) is self-similar, i.e., $\iota_{X_0} F = 0$ and $(\mathbf{D}_{X_0} + \frac{1}{\rho})\phi = 0$, where $X_0 = \partial_\rho$ in the hyperbolic coordinates (ρ, y, Θ). Then we have

$$\int_{\mathcal{H}_\rho} \frac{1}{2}\left(\frac{\cosh y}{\rho^2}|\phi|^2 + 2\frac{\sinh y}{\rho^2}\mathrm{Re}(\phi\overline{\mathbf{D}_y\phi}) + \cosh y\left(|\mathbf{D}\phi|^2_{\mathcal{H}_\rho} + |F|^2_{\mathcal{H}_\rho}\right)\right) \leq E, \tag{5.2}$$

where $|\mathbf{D}\phi|^2_{\mathcal{H}_\rho}, |F|^2_{\mathcal{H}_\rho}$ are to be defined in (5.27).

The next statement concerns the quantities

$$\mathcal{F}_{\partial C_{[t_0, t_1]}}[A, \phi] := \mathcal{E}_{S_{t_1}}[A, \phi] - \mathcal{E}_{S_{t_0}}[A, \phi], \quad \mathcal{G}_{\partial S_{t_1}}[\phi] := \frac{1}{t_1}\int_{\partial S_{t_1}}|\phi|^2. \tag{5.3}$$

Here, $\mathcal{F}_{\partial C_{[t_0, t_1]}}$ is the *energy flux* of (A, ϕ) through $\partial C_{[t_0, t_1]}$. For $\phi \in C_t(I; \dot{H}^1_x)$ and $t_1 \in I$, observe that $\mathcal{G}_{\partial S_{t_1}}[\phi]$ is well-defined by the trace theorem. In fact, $\phi\restriction_{\partial S_{t_1}} \in H^{1/2}(\partial S_{t_1})$.

Lemma 5.2 *Let (A, ϕ) be an admissible $C_t\mathcal{H}^1$ solution to* (MKG) *on $I \times \mathbb{R}^4$ where $I \subset \mathbb{R}^4$ is an open interval. Then for every $t_0, t_1 \in I$ with $t_0 \leq t_1$, the following statements hold:*

(1) *The energy flux on $\mathcal{F}_{\partial C_{[t_0, t_1]}}[A, \phi]$ is non-negative and additive, i.e.,*

$$\mathcal{F}_{\partial C_{[t_0, t_1]}}[A, \phi] = \mathcal{F}_{\partial C_{[t_0, t']}}[A, \phi] + \mathcal{F}_{\partial C_{[t', t_1]}}[A, \phi] \quad \text{for } t' \in [t_0, t_1]. \tag{5.4}$$

(2) *The following local Hardy's inequality holds on $\partial C_{[t_0, t_1]}$:*

$$\mathcal{G}_{\partial S_{t_0}}[\phi] + \int_{t_0}^{t_1} \mathcal{G}_{\partial S_t}[\phi]\frac{\mathrm{d}t}{t} \leq \mathcal{G}_{\partial S_{t_1}}[\phi] + \mathcal{F}_{\partial C_{[t_0, t_1]}}[A, \phi]. \tag{5.5}$$

Moreover, we also have

$$\mathcal{G}_{\partial S_{t_1}}[\phi] \leq \mathcal{E}_{(\{t\} \times \mathbb{R}^4) \setminus S_{t_1}}[A, \phi] \tag{5.6}$$

A consequence of Lemma 5.2 is a simple but crucial decay result for the two quantities defined in (5.3).

Corollary 5.3 *Let (A, ϕ) be an admissible $C_t \mathcal{H}^1$ solution to (MKG) on $I \times \mathbb{R}^4$ where $I \subset \mathbb{R}^4$ is an open interval. Then the following statements hold.*

(1) *If $(0, \delta] \subseteq I$ for some $\delta > 0$, then we have*

$$\lim_{t_1 \to 0} \mathcal{F}_{\partial C_{(0,t_1]}}[A, \phi] = 0, \quad \lim_{t_1 \to 0} \mathcal{G}_{\partial S_{t_1}}[\phi] = 0. \tag{5.7}$$

where $\mathcal{F}_{\partial C_{(0,t_1]}}[A, \phi] := \lim_{t_0 \to 0} \mathcal{F}_{\partial C_{[t_0,t_1]}}[A, \phi]$.
(2) *If $[\delta, \infty) \subseteq I$ for some $\delta > 0$, then we have*

$$\lim_{t_0, t_1 \to \infty} \mathcal{F}_{\partial C_{[t_0,t_1]}}[A, \phi] = 0, \quad \lim_{t_1 \to \infty} \mathcal{G}_{\partial S_{t_1}}[\phi] = 0. \tag{5.8}$$

The statements concerning $\mathcal{F}_{\partial C_{[t_0,t_1]}}$ follow from the monotonicity and boundedness of $\mathcal{E}_{S_t}$, whereas those concerning $\mathcal{G}_{\partial S_{t_1}}$ follow from (5.5), (5.6); we omit the straightforward details.

The decay statements (5.7) and (5.8) imply that the energy flux and the quantity $\mathcal{G}_{\partial S_t}[\phi]$ vanish as one approaches $(0, 0)$ or $t \to \infty$. In the ideal case when $\mathcal{F}_{\partial C_{[t_0,t_1]}} = 0$ and $\mathcal{G}_{\partial S_{t_1}} = 0$, the solution (A, ϕ) enjoys an additional monotonicity formula, namely

$$\int_{S_{t_1}} {}^{(X_0)} P_T[A, \phi] \, dx + \iint_{C_{[t_0,t_1]}} \frac{1}{\rho} |\iota_{X_0} F|^2$$

$$+ \frac{1}{\rho} \left| \left(\mathbf{D}_{X_0} + \frac{1}{\rho} \right) \phi \right|^2 \, dt dx = \int_{S_{t_0}} {}^{(X_0)} P_T[A, \phi] \, dx \tag{5.9}$$

where $X_0 = \partial_\rho$ in the hyperbolic coordinate system (ρ, y, Θ), $|\iota_{X_0} F|^2 :=$ $\mathbf{m}(\iota_{X_0} F, \iota_{X_0} F)$ (observe that $|\iota_{X_0} F|^2 \geq 0$) and ${}^{(X_0)} P_T[A, \phi]$ is to be defined below in Lemma 5.10. It turns out that the right-hand side is uniformly bounded by the conserved energy as $t_0 \to 0$, thereby breaking the scaling invariance. More precisely, the first term on the left-hand side precludes null concentration of energy, whereas the second term implies that rescalings of (A, ϕ) are asymptotically self-similar.

In application, however, the quantities $\mathcal{F}$ and $\mathcal{G}$ will be small but not necessarily zero. Hence we will rely on the following approximate version of (5.9) instead. Define

$$\rho_\varepsilon = \sqrt{(t+\varepsilon)^2 - r^2}, \quad X_\varepsilon = \rho_\varepsilon^{-1}((t+\varepsilon)\partial_t + r\partial_r),$$
$$|\iota_{X_\varepsilon} F|^2 := \mathbf{m}(\iota_{X_\varepsilon} F, \iota_{X_\varepsilon} F).$$

Proposition 5.4 *Let (A, ϕ) be an admissible $C_t\mathcal{H}^1$ solution to (MKG) on $[\varepsilon, 1] \times \mathbb{R}^4$, where $\varepsilon \in (0, 1)$. Suppose furthermore that (A, ϕ) satisfies*

$$\mathcal{E}_{S_1}[A, \phi] \le E, \quad \mathcal{F}_{\partial C_{[\varepsilon, 1]}}[A, \phi] \le \varepsilon^{\frac{1}{2}} E, \quad \mathcal{G}_{\partial S_1}[\phi] \le \varepsilon^{\frac{1}{2}} E. \quad (5.10)$$

Then

$$\int_{S_1} {}^{(X_\varepsilon)} P_T[A, \phi] \, dx + \iint_{C_{[\varepsilon, 1]}} \frac{1}{\rho_\varepsilon} |\iota_{X_\varepsilon} F|^2 + \frac{1}{\rho_\varepsilon} \left| \left(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon} \right) \phi \right|^2 \, dt dx \lesssim E$$
$$(5.11)$$

where the implicit constant is independent of ε, E. We refer to Lemma 5.10 for the computation of ${}^{(X_\varepsilon)} P_T[A, \phi]$.

Using Proposition 5.4, we can also establish a version of (5.9) that is localized away from the boundary of the cone. This statement will be useful for propagating lower bounds in a time-like region towards $(0, 0)$.

Proposition 5.5 *Let (A, ϕ) be an admissible $C_t\mathcal{H}^1$ solution to (MKG) on $[\varepsilon, 1] \times \mathbb{R}^4$, where $\varepsilon \in (0, 1)$. Suppose furthermore that (A, ϕ) satisfies (5.10). Then for $2\varepsilon \le \delta_0 < \delta_1 \le t_0 \le 1$, we have*

$$\int_{S_1^{\delta_1}} {}^{(X_0)} P_T[A, \phi] \, dx \le \int_{S_{t_0}^{\delta_0}} {}^{(X_0)} P_T[A, \phi] \, dx$$
$$+ C \left((\delta_1/t_0)^{\frac{1}{2}} + |\log(\delta_1/\delta_0)|^{-1} \right) E. \quad (5.12)$$

The rest of this section is devoted to the proofs of the above statements, and is organized as follows. In Sect. 5.1, we discuss ways of generating divergence identities for proving the above conservation laws and monotonicity formulae. We also introduce null decomposition, which will assist our computations below. In Sect. 5.2, we use them to prove (5.1) and Proposition 5.1. In Sect. 5.3, we introduce and prove a local version of Hardy's inequality and use it establish Lemma 5.2. Lastly, Sect. 5.4 is devoted to the proof of (5.9) and Propositions 5.4 and 5.5.

5.1 Divergence identities and null decomposition

The goal of this subsection is two-fold. First, we introduce methods for generating useful divergence identities for solutions to (MKG) that essentially arise

from Nöther's principle. Second, we define the notion of a null frame and the associated null decomposition of F and $\mathbf{D}\phi$, which will be useful for the computations below.

We first present the energy-momentum tensor formalism for generating divergence identities. This formalism is a way to exploit Nöther's principle (continuous symmetries lead to conserved quantities in a Lagrangian field theory) for external symmetries, i.e., symmetries of the base manifold $\mathbb{R}^{1+4}$ of (MKG). Let (A, ϕ) be a smooth solution to (MKG) on an open subset $\mathcal{O} \subseteq \mathbb{R}^{1+4}$. We define the energy-momentum tensor associated to (A, ϕ) as

$$\mathcal{Q}_{\mathbf{ab}}[A, \phi] = {}^{(\mathrm{M})}\mathcal{Q}_{\mathbf{ab}}[A]_{\mathbf{ab}} + {}^{(\mathrm{KG})}\mathcal{Q}_{\mathbf{ab}}[A, \phi] \tag{5.13}$$

where

$$^{(\mathrm{M})}\mathcal{Q}_{\mathbf{ab}}[A] = F_{\mathbf{a}}{}^{\mathbf{c}}F_{\mathbf{bc}} - \frac{1}{4}\mathbf{m}_{\mathbf{ab}}F_{\mathbf{cd}}F^{\mathbf{cd}} \tag{5.14}$$

$$^{(\mathrm{KG})}\mathcal{Q}_{\mathbf{ab}}[A, \phi] = \mathrm{Re}(\mathbf{D_a}\phi\overline{\mathbf{D_b}\phi}) - \frac{1}{2}\mathbf{m}_{\mathbf{ab}}\mathbf{D^c}\phi\overline{\mathbf{D_c}\phi} \tag{5.15}$$

Note that $\mathcal{Q}$ is a symmetric 2-tensor, which is gauge invariant at each point. Moreover, since (A, ϕ) is a smooth solution to (MKG), the energy-momentum tensor satisfies

$$\nabla^{\mathbf{a}}\mathcal{Q}_{\mathbf{ab}}[A, \phi] = 0. \tag{5.16}$$

Given a vector field X on $\mathcal{O}$, we define its deformation tensor to be the Lie derivative of the metric with respect to X, i.e., $^{(X)}\pi := \mathcal{L}_X\mathbf{m}$. Using covariant derivatives, $^{(X)}\pi$ also takes the form

$$^{(X)}\pi_{\mathbf{ab}} = \nabla_{\mathbf{a}}X_{\mathbf{b}} + \nabla_{\mathbf{a}}X_{\mathbf{b}}$$

We will denote the metric dual of $^{(X)}\pi$ by $^{(X)}\pi^{\sharp}$, i.e., $(^{(X)}\pi^{\sharp})^{\mathbf{ab}} = \mathbf{m^{ac}m^{bd}}{}^{(X)}\pi_{\mathbf{cd}}$. From its Lie derivative definition, the following formula for $^{(X)}\pi_{\mu\nu}$ in coordinates can be immediately derived:

$$^{(X)}\pi_{\mu\nu} = X(\mathbf{m}_{\mu\nu}) + \partial_{\mu}(X^{\alpha})\mathbf{m}_{\alpha\nu} + \partial_{\nu}(X^{\alpha})\mathbf{m}_{\alpha\mu} \tag{5.17}$$

Using the deformation tensor, we now define the associated 1- and 0-currents of (A, ϕ) as

$$^{(X)}J_{\mathbf{a}}[A,\phi] := \mathcal{Q}_{\mathbf{ab}}[A,\phi]X^{\mathbf{b}},$$

$$^{(X)}K[A,\phi] := \mathcal{Q}_{\mathbf{ab}}[A,\phi]\left(\frac{1}{2}{}^{(X)}\pi^{\sharp}\right)^{\mathbf{ab}}. \tag{5.18}$$

Then by (5.16) and the symmetry of $\mathcal{Q}[A,\phi]_{\mathbf{ab}}$, we obtain

$$\nabla^{\mathbf{a}}\left({}^{(X)}J_{\mathbf{a}}[A,\phi]\right) = {}^{(X)}K[A,\phi]. \tag{5.19}$$

Remark 5.6 Taking $X = T = \partial_t$ in the rectilinear coordinates $(t, x^1, \ldots, x^4)$, we have ${}^{(T)}\pi = 0$ (in other words, T is a Killing vector field) and hence ${}^{(T)}K = 0$. In fact, (5.19) is a local form of the standard conservation of energy (5.1). We refer to Sect. 5.2 for more details.

For a (smooth) scalar field ϕ satisfying the gauge covariant wave equation $\Box_A\phi = 0$, we introduce another way of generating divergence identities. This method corresponds to using Nöther's principle for the symmetry of the equation under the action of $\mathbb{C}$ viewed as the complexification of the gauge group $U(1)$. Given a $\mathbb{C}$-valued function w on an open subset of $\mathbb{R}^{1+4}$, we define its associated 1- and 0-currents by

$$^{(w)}J_{\mathbf{a}}[A,\phi] = (\mathrm{Re}\,w)\mathrm{Re}(\phi\overline{\mathbf{D_a}\phi}) - (\mathrm{Im}\,w)\mathrm{Im}(\phi\overline{\mathbf{D_a}\phi}) - \frac{1}{2}\nabla_{\mathbf{a}}(\mathrm{Re}\,w)|\phi|^2,$$

$$^{(w)}K[A,\phi] = (\mathrm{Re}\,w)\mathbf{D_a}\phi\overline{\mathbf{D^a}\phi} - \frac{1}{2}\Box(\mathrm{Re}\,w)|\phi|^2 - \nabla_{\mathbf{a}}(\mathrm{Im}\,w)\mathrm{Im}(\phi\overline{\mathbf{D^a}\phi}). \tag{5.20}$$

A simple computation[12] shows that the following conservation law holds:

$$\nabla^{\mathbf{a}}\left({}^{(w)}J_{\mathbf{a}}[A,\phi]\right) = {}^{(w)}K[A,\phi]. \tag{5.21}$$

Remark 5.7 Taking $w = -i$, we have

$$^{(w)}J_{\mathbf{a}} = \mathrm{Im}(\phi\overline{\mathbf{D_a}\phi}), \quad {}^{(w)}K = 0,$$

and (5.21) reduces to the well-known local conservation of charge.

Finally, we introduce the notion of a null frame and the associated null decomposition of $\mathbf{D}\phi$ and F, which are useful for computations concerning

---

[12] Alternatively, the identity below can be derived by multiplying the covariant wave equation for ϕ by $\overline{w\phi}$, taking the real part and differentiating by parts.

the energy-momentum tensor. At each point $p = (t_0, x_0) \in \mathbb{R}^{1+4}$, consider orthonormal vectors $\{e_a\}_{a=1,...,3}$ which are orthogonal to L and $\underline{L}$. Observe that each e_a is tangent to the sphere $\partial B_{t_0, r_0} := \{t_0\} \times \partial B_{r_0}(0)$ where $r_0 = |x_0|$. The set of vectors $\{L, \underline{L}, e_1, e_2, e_3\}$ at p is called a null frame at p associated to $L, \underline{L}$.

The $\mathbb{C}$-valued 1-form $\mathbf{D}\phi$ can be decomposed with respect to the null frame $\{L, \underline{L}, e_a\}$ as $\mathbf{D}_L\phi$, $\mathbf{D}_{\underline{L}}\phi$ and $\mathbf{D}_a\phi := \mathbf{D}_{e_a}\phi$, which is the null decomposition of $\mathbf{D}\phi$. A simple computation shows that

$$^{(KG)}\mathcal{Q}[A, \phi](L, L) = |\mathbf{D}_L\phi|^2,$$
$$^{(KG)}\mathcal{Q}[A, \phi](\underline{L}, \underline{L}) = |\mathbf{D}_{\underline{L}}\phi|^2, \quad ^{(KG)}\mathcal{Q}[A, \phi](L, \underline{L}) = |\mathbf{D}\!\!\!/\,\phi|^2 \quad (5.22)$$

where $|\mathbf{D}\!\!\!/\,\phi|^2 := \sum_{a=1,...,3} |\mathbf{D}\!\!\!/_a\phi|^2$.

Next, we define the null decomposition of the 2-form F with respect to $\{L, \underline{L}, e_a\}$ as

$$\alpha_a := F(L, e_a), \quad \underline{\alpha}_a := F(\underline{L}, e_a), \quad \varrho := \frac{1}{2}F(L, \underline{L}), \quad \sigma_{ab} := F(e_a, e_b).$$

Note that ϱ is a function, α_a, $\underline{\alpha}_b$ are 1-forms on $\partial B_{t_0, r_0}$ and σ_{ab} is a 2-form on $\partial B_{t_0, r_0}$. We define their pointwise absolute values as

$$|\alpha|^2 := \sum_{a=1,...,3} \alpha_a^2, \quad |\underline{\alpha}|^2 := \sum_{a=1,...,3} \underline{\alpha}_a^2, \quad |\sigma|^2 := \sum_{1 \le a < b \le 3} \sigma_{ab}^2.$$

This decomposition leads to the following simple formulae for the $L, \underline{L}$ components of $^{(M)}\mathcal{Q}$:

$$^{(M)}\mathcal{Q}[A](L, L) = |\alpha|^2, \quad ^{(M)}\mathcal{Q}[A](\underline{L}, \underline{L}) = |\underline{\alpha}|^2,$$
$$^{(M)}\mathcal{Q}[A](L, \underline{L}) = |\varrho|^2 + |\sigma|^2. \quad (5.23)$$

5.2 The standard energy identity and proof of Proposition 5.1

Consider the vector field T, which is equal to the coordinate vector field ∂_t in the rectilinear coordinates $(t, x^1, \ldots, x^4)$. It can be easily checked that T is Killing, i.e., $^{(T)}\pi = 0$. Contracting T with the energy-momentum tensor $\mathcal{Q}[A, \phi]$, we then obtain the local conservation of energy, i.e., given a smooth solution (A, ϕ) to (MKG) on an open subset $\mathcal{O} \subseteq \mathbb{R}^{1+4}$, we have

$$\nabla^a(^{(T)}J_a[A, \phi]) = 0 \quad \text{on } \mathcal{O}. \quad (5.24)$$

Since $T = \frac{1}{2}(L + \underline{L})$, we have

$$
\begin{aligned}
{}^{(T)}J_L[A, \phi] = \mathcal{Q}[A, \phi](T, L) &= \frac{1}{2}\left(|\mathbf{D}_L\phi|^2 + |\mathbf{D}\!\!\!/\,\phi|^2\right) \\
&+ \frac{1}{2}\left(|\alpha|^2 + |\varrho|^2 + |\sigma|^2\right),
\end{aligned} \tag{5.25}
$$

$$
\begin{aligned}
{}^{(T)}J_{\underline{L}}[A, \phi] = \mathcal{Q}[A, \phi](T, \underline{L}) &= \frac{1}{2}\left(|\mathbf{D}_{\underline{L}}\phi|^2 + |\mathbf{D}\!\!\!/\,\phi|^2\right) \\
&+ \frac{1}{2}\left(|\underline{\alpha}|^2 + |\varrho|^2 + |\sigma|^2\right).
\end{aligned} \tag{5.26}
$$

Given a (measurable) subset $S \subseteq \{t\} \times \mathbb{R}^4$ for some $t \in \mathbb{R}$, the above computation implies

$$
\mathcal{E}_S[A, \phi] = \int_S {}^{(T)}J_T[A, \phi]\, dx.
$$

We are now ready to give a quick proof of (5.1). For a classical solution (A, ϕ) in the class $C_t\mathcal{H}^1([t_0, t_1] \times \mathbb{R}^4)$, the standard energy conservation (5.1) follows by integrating (5.24) over $(t_0, t_1) \times \mathbb{R}^4$ and applying the divergence theorem. The case of an admissible solution then easily follows by approximation.

We conclude this subsection with a proof of Proposition 5.1.

Proof of Proposition 5.1 Note that $X_0 = \partial_\rho$ and $T = \cosh y\partial_\rho - \sinh y(\rho^{-1}\partial_y)$ in the hyperbolic coordinates (ρ, y, Θ). In the following computation, we use the orthonormal frame $\{\partial_\rho, \rho^{-1}\partial_y, e_{\mathfrak{a}}\}$ at each point, where $\{e_{\mathfrak{a}}\}_{\mathfrak{a}=1,2,3}$ is an orthonormal frame tangent to the constant ρ, y sphere as before. Then we compute

$$
\begin{aligned}
{}^{(KG)}\mathcal{Q}[A, \phi](\partial_\rho, \partial_\rho) &= \frac{1}{2}\left(|\mathbf{D}_\rho\phi|^2 + |\rho^{-1}\mathbf{D}_y\phi|^2 + |\mathbf{D}\!\!\!/\,\phi|^2\right) \\
{}^{(KG)}\mathcal{Q}[A, \phi](\partial_\rho, \rho^{-1}\partial_y) &= \mathrm{Re}(\mathbf{D}_\rho\phi\overline{\rho^{-1}\mathbf{D}_y\phi}), \\
{}^{(M)}\mathcal{Q}[A, \phi](\partial_\rho, \partial_\rho) &= \frac{1}{2}F(\partial_\rho, \rho^{-1}\partial_y)^2 + \frac{1}{2}\sum_{\mathfrak{a}=1,\ldots,3}F(\partial_\rho, e_{\mathfrak{a}})^2 \\
&+ \frac{1}{2}\sum_{\mathfrak{a}=1,\ldots,3}\rho^{-2}F(\partial_y, e_{\mathfrak{a}})^2 \\
&+ \frac{1}{2}\sum_{1\leq\mathfrak{a}<\mathfrak{b}\leq 3}F(e_{\mathfrak{a}}, e_{\mathfrak{b}})^2, \\
{}^{(M)}\mathcal{Q}[A, \phi](\partial_\rho, \rho^{-1}\partial_y) &= \sum_{\mathfrak{a}=1,\ldots,3}F(\partial_\rho, e_{\mathfrak{a}})F(\rho^{-1}\partial_y, e_{\mathfrak{a}}).
\end{aligned}
$$

Fig. 1 Domain of
integration for the proof of
Proposition 5.1
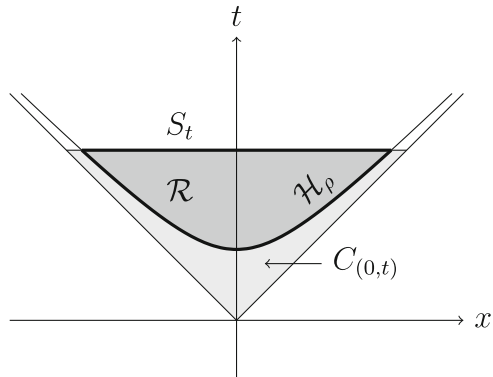
By the self-similarity conditions $\iota_{\partial_\rho} F = F(\partial_\rho, \cdot) = 0$ and $(\mathbf{D}_\rho + \frac{1}{\rho})\phi = 0$,
we have

$$^{(T)}J_\rho[A, \phi] = \cosh y \, \mathcal{Q}[A, \phi](\partial_\rho, \partial_\rho) - \sinh y \, \mathcal{Q}[A, \phi](\rho^{-1}\partial_y, \partial_\rho)$$

$$= \frac{1}{2}\left(\frac{\cosh y}{\rho^2}|\phi|^2 + 2\frac{\sinh y}{\rho^2}\mathrm{Re}(\phi\overline{\mathbf{D}_y\phi})\right.$$

$$\left. + \cosh y \left(|\mathbf{D}\phi|^2_{\mathcal{H}_\rho} + |F|^2_{\mathcal{H}_\rho}\right)\right)$$

where

$$|\mathbf{D}\phi|^2_{\mathcal{H}_\rho} := \left(g_{\mathcal{H}_\rho}^{-1}\right)^{\mathbf{ab}} \mathbf{D}_{\mathbf{a}}\phi\overline{\mathbf{D}_{\mathbf{b}}\phi}, \ |F|^2_{\mathcal{H}_\rho} := \frac{1}{2}\left(g_{\mathcal{H}_\rho}^{-1}\right)^{\mathbf{ac}}\left(g_{\mathcal{H}_\rho}^{-1}\right)^{\mathbf{bd}} F_{\mathbf{ab}}F_{\mathbf{cd}}, \quad (5.27)$$

and $g_{\mathcal{H}_\rho}^{-1} = \rho^{-2}\partial_y \cdot \partial_y + \sum_{\mathfrak{a}=1,2,3} e_{\mathfrak{a}} \cdot e_{\mathfrak{a}}$ is the induced metric on $\mathcal{H}_\rho$.

We are ready to complete the proof. Denote by $\mathcal{H}_{>\rho}$ the region $\{(\rho', y', \Theta') : \rho' > \rho\}$. Integrate (5.24) over the region $\mathcal{R} := C_{(0,t)} \cap \mathcal{H}_{>\rho}$, whose boundary is $(S_t \cap \mathcal{H}_{>\rho}) \cup (\mathcal{H}_\rho \cap C_{(0,t)})$, and apply the divergence theorem; see Fig. 1. Then taking $t \to \infty$, the desired estimate (5.2) on $\mathcal{H}_\rho$ follows. $\square$

5.3 A localized Hardy's inequality and proof of Lemma 5.2

We begin by stating a very general identity (valid for any dimension $d \geq 3$), which can be thought of as Hardy's inequality with all the errors terms explicit.

Lemma 5.8 *Let ϕ be a smooth $\mathbb{C}$-valued function and A be a smooth 1-form on $\mathbb{R}^d$ ($d \geq 3$). Then for $0 < r_1 < r_2$, we have*

$$\int_{r_1}^{r_2} \int \frac{1}{r^2} |\phi|^2 r^{d-1} \, d\sigma_{\mathbb{S}^{d-1}} \, dr$$

$$+ \int_{r_1}^{r_2} \int \left| \frac{2}{d-2} \mathbf{D}_r \phi + \frac{1}{r} \phi \right|^2 r^{d-1} \, d\sigma_{\mathbb{S}^{d-1}} \, dr$$

$$= \left(\frac{2}{d-2} \right)^2 \int_{r_1}^{r_2} \int |\mathbf{D}_r \phi|^2 r^{d-1} \, d\sigma_{\mathbb{S}^{d-1}} \, dr$$

$$+ \frac{2}{d-2} \int |\phi|^2 r^{d-2} \, d\sigma_{\mathbb{S}^{d-1}} \Big|_{r=r_1}^{r_2}. \tag{5.28}$$

We omit the proof, which is a simple algebra plus an application of the fundamental theorem of calculus in r. Specializing to $d = 4$ and rearranging some terms, we obtain

$$\int_{\{r=r_1\}} \frac{|\phi|^2}{r} r^3 d\sigma_{\mathbb{S}^3} + \int_{r_1}^{r_2} \int \frac{1}{r^2} |\phi|^2 r^3 \, d\sigma_{\mathbb{S}^3} \, dr$$

$$+ \int_{r_1}^{r_2} \int |r^{-1} \mathbf{D}_r (r\phi)|^2 r^3 \, d\sigma_{\mathbb{S}^3} \, dr$$

$$= \int_{\{r=r_2\}} \frac{|\phi|^2}{r} r^3 d\sigma_{\mathbb{S}^3} + \int_{r_1}^{r_2} \int |\mathbf{D}_r \phi|^2 r^3 \, d\sigma_{\mathbb{S}^3} \, dr. \tag{5.29}$$

The last term on the left-hand side of (5.29) is always non-negative; moreover, for $\phi \in \mathcal{S}(\mathbb{R}^4)$, the first term on the right-hand side vanishes as $r_2 \to \infty$. By approximation, the following gauge invariant version of Hardy's inequality on $\mathbb{R}^4$ follows.

Corollary 5.9 *Let* $\phi, A \in \dot{H}^1(\mathbb{R}^4) \cap L^4(\mathbb{R}^4)$. *Then* $r^{-1}\phi \in L^2(\mathbb{R}^4)$ *and* $\phi \upharpoonright_{\partial B_r} \in L^2(\partial B_r)$ *for every* $r > 0$. *Moreover, we have*

$$\left\| \frac{\phi}{r} \right\|_{L^2(\mathbb{R}^4)}^2 + \sup_{r>0} \frac{1}{r} \|\phi\|_{L^2(\partial B_r)}^2 \leq \|\mathbf{D}_r \phi\|_{L^2(\mathbb{R}^4)}^2. \tag{5.30}$$

We are ready to establish Lemma 5.2.

Proof of Lemma 5.2 We first consider the case when (A, ϕ) is smooth. Then by local conservation of energy, we have

$$\mathcal{F}_{\partial C_{[t_0, t_1]}} = \frac{1}{2} \int_{\partial C_{[t_0, t_1]}} {}^{(T)} J_L[A, \phi] r^3 \, dv d\sigma_{\mathbb{S}^3}$$

and hence the non-negativity and additivity are obvious. The first local Hardy's inequality (5.5) is a consequence of (5.29) applied to the hypersurface

$\partial C_{[t_0,t_1]} = \{u = 0, \ r \in [t_0, t_1]\}$ in the coordinate system (u, r, Θ), whereas the second local Hardy's inequality (5.6) follows from a similar argument used to derive Corollary 5.9.

Now we turn to the general case. Since (A, ϕ) is an admissible $C_t \mathcal{H}^1$ solution, there exists a sequence of smooth solutions converging to (A, ϕ) in $C_t \mathcal{H}^1(I \times \mathbb{R}^4)$. Since all quantities in the conclusions of the lemma are continuous with respect to the $C_t \mathcal{H}^1(I \times \mathbb{R}^4)$ topology, the general case follows from the smooth case by approximation. $\square$

5.4 Monotonicity formulae and proofs of Propositions 5.4, 5.5

Here we derive monotonicity formulae associated with the vector fields X_ε, which are defined in the polar coordinates as

$$X_\varepsilon = \frac{1}{\rho_\varepsilon}((t + \varepsilon)\partial_t + r \partial_r), \quad \rho_\varepsilon = \sqrt{(t + \varepsilon)^2 - r^2}, \qquad (5.31)$$

where $\varepsilon \geq 0, t > -\varepsilon$.

The starting point for derivation of the monotonicity formula (5.9), as well as Propositions 5.4 and 5.5, is to contract the energy-momentum tensor $\mathcal{Q}$ with one of the vector fields X_ε. Due to the unfavorable contribution of $^{(KG)}\mathcal{Q}$, however, several additional modifications are necessary. To simplify the discussion, we first restrict to the case $\varepsilon = 0$. The reader should keep in mind that the general case follows simply by translating in time by ε.

Using the formula (5.17), we compute

$$\frac{1}{2}{}^{(X_0)}\pi^\sharp = \frac{1}{\rho^3}\left(\partial_y \cdot \partial_y + \frac{1}{\sinh^2 y}\left(g_{\mathbb{S}^3}^{-1}\right)\right) = \frac{1}{\rho}\left(\mathbf{m}^{-1} + X_0 \cdot X_0\right).$$

Hence we have

$$\begin{aligned}
{}^{(X_0)}K &= {}^{(M)}\mathcal{Q}_{\mathbf{ab}}\left(\frac{1}{2}{}^{(X_0)}\pi^\sharp\right)^{\mathbf{ab}} + {}^{(KG)}\mathcal{Q}_{\mathbf{ab}}\left(\frac{1}{2}{}^{(X_0)}\pi^\sharp\right)^{\mathbf{ab}} \\
&= \frac{1}{\rho}|\iota_{X_0}F|^2 + \frac{1}{\rho}|\mathbf{D}_{X_0}\phi|^2 - \frac{1}{\rho}\mathbf{D_a}\phi\overline{\mathbf{D^a}\phi}.
\end{aligned} \qquad (5.32)$$

where $|\iota_{X_0}F|^2 = \mathbf{m}(\iota_{X_0}F, \iota_{X_0}F) \geq 0$, since X_0 is time-like. The first term on (5.32) is satisfactory in view of our goal (5.9), but the rest is not. To remove the last term, we use the currents $^{(w_0)}J$ and $^{(w_0)}K$ with $w_0 = \frac{1}{\rho}$ and compute

$$^{(X_0)}K + {}^{(w_0)}K = \frac{1}{\rho}|\iota_{X_0}F|^2 + \frac{1}{\rho}|\mathbf{D}_{X_0}\phi|^2 - \frac{1}{\rho^3}|\phi|^2. \qquad (5.33)$$

Now we introduce an auxiliary divergence identity, which is related to Hardy's inequality in the ρ variable. Define $^{(\mathcal{H}_0)}J[\phi]$ in the hyperbolic coordinates (ρ, y, Θ) by

$$^{(\mathcal{H}_0)}J_\rho[\phi] := -\frac{|\phi|^2}{\rho^2}, \tag{5.34}$$

where the remaining components are set to be zero. Define also

$$^{(\mathcal{H}_0)}K[\phi] := \frac{2}{\rho^3}|\phi|^2 + \frac{1}{\rho^2}\partial_\rho|\phi|^2. \tag{5.35}$$

Then a simple computation shows that

$$\nabla^{\mathbf{a}}(^{(\mathcal{H}_0)}J_{\mathbf{a}}[\phi]) = {}^{(\mathcal{H}_0)}K[\phi]. \tag{5.36}$$

Since $\partial_\rho|\phi|^2 = 2\mathrm{Re}(\phi\overline{\mathbf{D}_\rho\phi})$ and $X_0 = \partial_\rho$, we arrive at

$$^{(X_0)}K + {}^{(w_0)}K + {}^{(\mathcal{H}_0)}K = \frac{1}{\rho}|\iota_{X_0}F|^2 + \frac{1}{\rho}\left|\left(\mathbf{D}_{X_0} + \frac{1}{\rho}\right)\phi\right|^2, \tag{5.37}$$

which is precisely the integrand in the space-time integral in (5.9).

The preceding computation suggests that we should define a new 1- and 0-currents by $^{(X_0)}J + {}^{(w_0)}J + {}^{(\mathcal{H}_0)}J$ and $^{(X_0)}K + {}^{(w_0)}K + {}^{(\mathcal{H}_0)}K$, respectively. To make the L and $\underline{L}$ components of the 1-current look more favorable, however, it turns out to be convenient to add in an auxiliary current $^{(\mathcal{N}_0)}J$ defined by

$$\begin{aligned}
^{(\mathcal{N}_0)}J_L[\phi] &= \frac{1}{2r^3}L\left(r^3\frac{t}{\rho r}|\phi|^2\right), \\
^{(\mathcal{N}_0)}J_{\underline{L}}[\phi] &= -\frac{1}{2r^3}\underline{L}\left(r^3\frac{t}{\rho r}|\phi|^2\right),
\end{aligned} \tag{5.38}$$

where the remaining components are set to be zero. By equality of mixed partials $L\underline{L} = 4\partial_v\partial_u = 4\partial_u\partial_v = \underline{L}L$, it follows that

$$\nabla^{\mathbf{a}}\left(^{(\mathcal{N}_0)}J_{\mathbf{a}}[\phi]\right) = 0. \tag{5.39}$$

For $^{(X_0)}P := {}^{(X_0)}J + {}^{(w_0)}J + {}^{(\mathcal{H}_0)}J + {}^{(\mathcal{N}_0)}J$, we claim that

$$\begin{aligned}
^{(X_0)}P_L &= \frac{1}{2}\left(\frac{v}{u}\right)^{\frac{1}{2}}\left(|r^{-1}\mathbf{D}_L(r\phi)|^2 + |\alpha|^2\right) \\
&\quad + \frac{1}{2}\left(\frac{u}{v}\right)^{\frac{1}{2}}\left(|\mathbf{\not{D}}\phi|^2 + \frac{|\phi|^2}{r^2} + |\varrho|^2 + |\sigma|^2\right),
\end{aligned} \tag{5.40}$$

$$^{(X_0)}P_{\underline{L}} = \frac{1}{2}\left(\frac{u}{v}\right)^{\frac{1}{2}}\left(|r^{-1}\mathbf{D}_{\underline{L}}(r\phi)|^2 + |\underline{\alpha}|^2\right)$$
$$+ \frac{1}{2}\left(\frac{v}{u}\right)^{\frac{1}{2}}\left(|\mathbf{D}\!\!\!/\phi|^2 + \frac{|\phi|^2}{r^2} + |\varrho|^2 + |\sigma|^2\right). \qquad (5.41)$$

We will prove (5.40), leaving the task of verifying (5.41) to the reader. Using the relations

$$\rho^2 = uv, \quad X_0 = \frac{1}{2}\left(\frac{v}{\rho}L + \frac{u}{\rho}\underline{L}\right),$$

and the null decomposition formulae (5.22), (5.23), we have

$$^{(X_0)}J_L[A,\phi] = \frac{1}{2}\left(\frac{v}{\rho}|\mathbf{D}_L\phi|^2 + \frac{u}{\rho}|\mathbf{D}\!\!\!/\phi|^2\right) + \frac{1}{2}\left(\frac{v}{\rho}|\alpha|^2 + \frac{u}{\rho}\left(|\varrho|^2 + |\sigma|^2\right)\right).$$

On the other hand, we compute

$$^{(w_0)}J_L[A,\phi] = \frac{1}{\rho}\mathrm{Re}(\phi\overline{\mathbf{D}_L\phi}) + \frac{1}{2}\frac{1}{\rho v}|\phi|^2, \quad ^{(\mathcal{H}_0)}J_L[\phi] = -\frac{1}{\rho v}|\phi|^2.$$

To prove (5.40), it suffices to verify

$$\frac{1}{2}\frac{v}{\rho}|\mathbf{D}_L\phi|^2 + {}^{(w_0)}J_L[A,\phi] + {}^{(\mathcal{H}_0)}J_L[\phi] + {}^{(\mathcal{N}_0)}J_L[\phi]$$
$$= \frac{1}{2}\frac{v}{\rho}|r^{-1}\mathbf{D}_L(r\phi)|^2 + \frac{1}{2}\frac{u}{\rho}\frac{|\phi|^2}{r^2}. \qquad (5.42)$$

For this purpose, it is convenient to work with $\psi = r\phi$. We have

$$\text{LHS of } (5.42) = \frac{1}{2}\frac{v}{\rho}|\mathbf{D}_L(\psi/r)|^2 + \frac{1}{\rho r}\mathrm{Re}(\psi\overline{\mathbf{D}_L(\psi/r)})$$
$$+ \frac{1}{2}\frac{1}{\rho v}\frac{|\psi|^2}{r^2} - \frac{1}{\rho v}\frac{|\psi|}{r^2} + \frac{1}{2r^3}L\left(\frac{t}{\rho}|\psi|^2\right)$$
$$= \frac{1}{2}\frac{v}{\rho}|r^{-1}\mathbf{D}_L\psi|^2 + \frac{1}{2}\left(\frac{v}{\rho r^2} - \frac{2}{\rho r} - \frac{1}{\rho v} + \frac{1}{r}L(t/\rho)\right)\frac{|\psi|^2}{r^2}$$

Since $r^{-1}L(t/\rho) = 1/(\rho r) - t/(\rho r v) = 1/(\rho v)$, we see that

$$\frac{v}{\rho r^2} - \frac{2}{\rho r} - \frac{1}{\rho v} + \frac{1}{r}L(t/\rho) = \frac{v}{\rho r^2} - \frac{2}{\rho r} = \frac{u}{\rho r^2},$$

which establishes (5.42), and hence (5.40).

We now return to the general case $\varepsilon \geq 0$. Define $^{(X_\varepsilon)}J$, $^{(w_\varepsilon)}J$, $^{(\mathcal{H}_\varepsilon)}J$, $^{(\mathcal{N}_\varepsilon)}J$ and their 0-current counterparts by pulling back the $\varepsilon = 0$ versions defined above along the map $(t, r, \Theta) \mapsto (t + \varepsilon, r, \Theta)$. For $^{(X_\varepsilon)}J$, $^{(w_\varepsilon)}J$, $^{(X_\varepsilon)}K$ and $^{(w_\varepsilon)}K$, note that this definition agrees with that from Sect. 5.1 using X_ε as in (5.31) and $w_\varepsilon := 1/\rho_\varepsilon$. Let

$$
^{(X_\varepsilon)}P[A, \phi] := {}^{(X_\varepsilon)}J[A, \phi] + {}^{(w_\varepsilon)}J[A, \phi] + {}^{(\mathcal{H}_\varepsilon)}J[\phi] + {}^{(\mathcal{N}_\varepsilon)}J[\phi],
$$
$$
^{(X_\varepsilon)}Q[A, \phi] := {}^{(X_\varepsilon)}K[A, \phi] + {}^{(w_\varepsilon)}K[A, \phi] + {}^{(\mathcal{H}_\varepsilon)}K[\phi]. \tag{5.43}
$$

We summarize the discussion so far in the following lemma, which follows easily by pulling back the above computations along $(t, r, \Theta) \mapsto (t + \varepsilon, r, \Theta)$.

Lemma 5.10 *Let (A, ϕ) be a smooth solution to (MKG) on an open subset $\mathcal{O} \subseteq C_{(0,\infty)}$. The 1- and 0-currents $^{(X_\varepsilon)}P[A, \phi]$ and $^{(X_\varepsilon)}Q$ obeys the divergence identity*

$$
\nabla^{\mathbf{a}} \left({}^{(X_\varepsilon)}P_{\mathbf{a}}[A, \phi] \right) = {}^{(X_\varepsilon)}Q[A, \phi], \tag{5.44}
$$

where $^{(X_\varepsilon)}Q = {}^{(X_\varepsilon)}Q[A, \phi]$ takes the form

$$
^{(X_\varepsilon)}Q = \frac{1}{\rho_\varepsilon} |\iota_{X_\varepsilon} F|^2 + \frac{1}{\rho_\varepsilon} \left| \left(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon} \right) \phi \right|^2. \tag{5.45}
$$

Here, $|\iota_{X_\varepsilon} F|^2 = \mathbf{m}(\iota_{X_\varepsilon} F, \iota_{X_\varepsilon} F) \geq 0$. Moreover, the L and $\underline{L}$ components of $^{(X_\varepsilon)}P = {}^{(X_\varepsilon)}P[A, \phi]$ take the form

$$
^{(X_\varepsilon)}P_L = \frac{1}{2} \left(\frac{v_\varepsilon}{u_\varepsilon} \right)^{\frac{1}{2}} \left(|r^{-1}\mathbf{D}_L(r\phi)|^2 + |\alpha|^2 \right)
$$
$$
+ \frac{1}{2} \left(\frac{u_\varepsilon}{v_\varepsilon} \right)^{\frac{1}{2}} \left(|\not{\mathbf{D}}\phi|^2 + \frac{|\phi|^2}{r^2} + |\varrho|^2 + |\sigma|^2 \right), \tag{5.46}
$$
$$
^{(X_\varepsilon)}P_{\underline{L}} = \frac{1}{2} \left(\frac{u_\varepsilon}{v_\varepsilon} \right)^{\frac{1}{2}} \left(|r^{-1}\mathbf{D}_{\underline{L}}(r\phi)|^2 + |\underline{\alpha}|^2 \right)
$$
$$
+ \frac{1}{2} \left(\frac{v_\varepsilon}{u_\varepsilon} \right)^{\frac{1}{2}} \left(|\not{\mathbf{D}}\phi|^2 + \frac{|\phi|^2}{r^2} + |\varrho|^2 + |\sigma|^2 \right), \tag{5.47}
$$

where $v_\varepsilon := (t + \varepsilon) + r$ and $u_\varepsilon := (t + \varepsilon) - r$.

Here we give a quick proof of (5.9) for a smooth solution (A, ϕ) on $\mathbb{R}^{1+4}$. By $\mathcal{F}_{\partial C_{[t_0,t_1]}} = 0$, $\mathcal{G}_{\partial S_{t_1}} = 0$ and Lemma 5.2, note that ϕ and the tangential components of F (i.e., α, ϱ, σ) vanish on the boundary $\partial C_{[t_0,t_1]}$. Integrate

(5.44) with $\varepsilon = 0$ over $C_{[t_0,t_1]}$ and apply the divergence theorem. The boundary term on $\partial C_{[t_0,t_1]}$ vanishes thanks to $\phi, \alpha, \varrho, \sigma = 0$, and thus (5.9) follows.

In the preceding proof, however, note from (5.40) that there is a weight $(\frac{v}{u})^{1/2}$ in the boundary term, which would blow up if $\mathbf{D}_L(r\phi)$ and α_A were not exactly zero on $\partial C_{[t_0,t_1]}$. We now turn to the proof of Proposition 5.4, whose goal is exactly to deal with this issue.

Proof of Proposition 5.4 As the hypothesis (5.10) and the conclusion (5.11) only involve quantities which are continuous with respect to the $C_t\mathcal{H}^1(I \times \mathbb{R}^4)$ topology, it suffices to consider the case when (A, ϕ) is smooth. Integrating (5.44) with $\varepsilon > 0$ over $C_{[\varepsilon,1]}$ and integrating by parts, we obtain

$$\int_{S_1} {}^{(X_\varepsilon)}P_T[A, \phi]\,\mathrm{d}x + \iint_{C_{[\varepsilon,1]}} \frac{1}{\rho_\varepsilon}|\iota_{X_\varepsilon}F|^2 + \frac{1}{\rho_\varepsilon}\left|\left(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon}\right)\phi\right|^2 \mathrm{d}t\,\mathrm{d}x$$
$$= \int_{S_\varepsilon} {}^{(X_\varepsilon)}P_T[A, \phi]\,\mathrm{d}x + \frac{1}{2}\int_{\partial C_{[\varepsilon,1]}} {}^{(X_\varepsilon)}P_L[A, \phi]r^3\,\mathrm{d}v\mathrm{d}\sigma_{\mathbb{S}^3}. \qquad (5.48)$$

We claim that the right-hand side is bounded from above by $\lesssim E$. We begin with the first term. On S_ε, we have the pointwise bound

$${}^{(X_\varepsilon)}P_T[A, \phi] \lesssim {}^{(T)}P_T[A, \phi] + \frac{1}{r^2}|\phi|^2,$$

since $u_\varepsilon/v_\varepsilon \simeq 1$ and $v_\varepsilon/u_\varepsilon \simeq 1$ on S_ε. By (5.10), Lemma 5.2 and (5.29) applied to ϕ on S_ε with $r_1 = 0, r_2 = \varepsilon$, it follows that the first term on the right-hand side of (5.48) is bounded by $\lesssim E$.

We now consider the last term in (5.48). On $\partial C_{[\varepsilon,1]}$, we have

$${}^{(X_\varepsilon)}P_L[A, \phi] \lesssim \varepsilon^{-\frac{1}{2}}\left(|\mathbf{D}_L\phi|^2 + \frac{1}{r^2}|\phi|^2 + |\alpha|^2\right) + {}^{(T)}J_L[A, \phi],$$

Then by (5.10), Lemma 5.2 and the fact that $t = r$ on ∂C, the last term in (5.48) is bounded by $\lesssim E$ as desired.

We end this section with a proof of Proposition 5.5.

Proof of Proposition 5.5 As before, by approximation, it suffices to consider the case when (A, ϕ) is smooth. Let $\delta \in [\delta_0, \delta_1]$ be a number to be determined below. Integrating (5.44) with $\varepsilon = 0$ over $C^\delta_{[t_0,1]}$ and using the divergence theorem, we see that (5.12) would follow if there exists $\delta \in [\delta_0, \delta_1]$ such that

$$\int_{\partial C^\delta_{[t_0,1]}} {}^{(X_0)}P_L[A, \phi]r^3\,\mathrm{d}v\mathrm{d}\sigma_{\mathbb{S}^3} \lesssim \left((\delta_1/t_0)^{\frac{1}{2}} + |\log(\delta_1/\delta_0)|^{-1}\right)E. \qquad (5.49)$$

The contribution of the term with the weight $(u_0/v_0)^{1/2}$ in (5.46) is easy to treat; indeed, using localized Hardy's inequality and local conservation of energy, we have

$$\int_{\partial C^\delta_{[t_0,1]}} \frac{1}{2} \left(\frac{u}{v}\right)^{\frac{1}{2}} \left(|\mathbf{D}\phi|^2 + \frac{|\phi|^2}{r^2} + |\varrho|^2 + |\sigma|^2\right) r^3 \, dv d\sigma_{\mathbb{S}^3}$$

$$\lesssim \left(\frac{\delta_1}{t_0}\right)^{1/2} \left(\int_{\partial C^\delta_{[t_0,1]}} {}^{(T)}J_L[A,\phi] \, r^3 \, dv d\sigma_{\mathbb{S}^3}\right.$$

$$\left. + \mathcal{E}_{S_1 \setminus S_1^\delta}[A,\phi] + \mathcal{G}_{S_1}[\phi]\right) \lesssim \left(\frac{\delta_1}{t_0}\right)^{1/2} E.$$

It remains to treat the term with the weight $(v_0/u_0)^{1/2}$ in (5.46). Note that

$$r^{-1}\mathbf{D}_L(r\phi) = \left(\mathbf{D}_L + \frac{1}{r}\right)\phi = 2\left(\frac{u_\varepsilon}{v_\varepsilon}\right)^{\frac{1}{2}} \left(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon}\right)\phi$$

$$- \left(\frac{u_\varepsilon}{v_\varepsilon}\right)\mathbf{D}_{\underline{L}}\phi + \left(\frac{u_\varepsilon}{v_\varepsilon}\right)\frac{1}{r}\phi,$$

$$\alpha_{\mathfrak{a}} = F(L, e_{\mathfrak{a}}) = 2\left(\frac{u_\varepsilon}{v_\varepsilon}\right)^{\frac{1}{2}} F(X_\varepsilon, e_{\mathfrak{a}}) - \left(\frac{u_\varepsilon}{v_\varepsilon}\right)F(\underline{L}, e_{\mathfrak{a}}).$$

Note that $u \leq u_\varepsilon$ and $v \leq v_\varepsilon$. Furthermore $u_\varepsilon \leq 2u$ on $\partial C^\delta_{[t_0,1]}$ since $2\varepsilon \leq \delta_0$. Hence,

$$\int_{\partial C^\delta_{[t_0,1]}} \frac{1}{2}\left(\frac{v}{u}\right)^{\frac{1}{2}} \left(|r^{-1}\mathbf{D}_L(r\phi)|^2 + |\alpha|^2\right) r^3 \, dv d\sigma_{\mathbb{S}^3}$$

$$\lesssim \int_{\partial C^\delta_{[t_0,1]}} \frac{u}{\rho_\varepsilon} \left(\left|\left(\mathbf{D}_{X_\varepsilon} + \frac{1}{\rho_\varepsilon}\right)\phi\right|^2 + |\iota_{X_\varepsilon}F|^2\right)$$

$$+ \frac{u^{\frac{3}{2}}}{v^{\frac{3}{2}}} \left(|\mathbf{D}_{\underline{L}}\phi|^2 + \frac{1}{r^2}|\phi|^2 + |\underline{\alpha}|^2\right) r^3 \, dv d\sigma_{\mathbb{S}^3}. \qquad (5.50)$$

We claim that the integral of the right-hand side over $\delta_0 \leq u \leq \delta_1$ with respect to $u^{-1}du$ is bounded by E. Then by the pigeonhole principle, there would exist $\delta \in [\delta_0, \delta_1]$ such that the left-hand side of (5.50) is bounded by $\lesssim |\log(\delta_1/\delta_0)|^{-1}E$, as desired.

For the contribution of the first term, the claim follows directly from Proposition 5.4. For the second term, we have

$$\iint_{C^{\delta_0}_{[t_0,1]} \setminus C^{\delta_1}_{[t_0,1]}} \frac{u^{\frac{1}{2}}}{v^{\frac{3}{2}}} \left(|\mathbf{D}_L \phi|^2 + \frac{1}{r^2} |\phi|^2 + |\underline{\alpha}|^2 \right) dt dx$$

$$\lesssim \iint_{C^{\delta_0}_{[t_0,1]} \setminus C^{\delta_1}_{[t_0,1]}} \frac{\delta_1^{\frac{1}{2}}}{t^{\frac{3}{2}}} {}^{(T)} J_T[A,\phi] \, dt dx \lesssim \left(\frac{\delta_1}{t_0} \right)^{\frac{1}{2}} E,$$

which is sufficient to prove the claim since $\delta_1 \leq t_0$.   □

6 Local strong compactness and weak solutions to (MKG)

The first goal of this section is to establish the following local strong compactness result for asymptotically stationary (see (6.2) below) sequences of solutions to (MKG) with small energy.

Proposition 6.1 *There exists a universal constant $\epsilon_0 > 0$ such that the following holds. Let $B = B_1(x_0) \subseteq \mathbb{R}^4$ be an open ball of unit radius centered at x_0, and let $(A^{(n)}, \phi^{(n)})$ be a sequence of admissible $C_t \mathcal{H}^1$ solutions to (MKG) in $(-2, 2) \times 8B$ such that*

$$\mathcal{E}_{\{0\} \times 8B} \left[A^{(n)}, \phi^{(n)} \right] + \|\phi^{(n)}(0, x)\|^2_{L^2_x(8B)} \leq \epsilon_0^2. \tag{6.1}$$

Suppose furthermore that $(A^{(n)}, \phi^{(n)})$ is asymptotically stationary in the sense that

$$\iint_{(-2,2) \times 2B} |\iota_X F^{(n)}|^2 + |(\mathbf{D}^{(n)}_X + b) \phi^{(n)}|^2 \, dt dx \to 0 \quad as \ n \to \infty, \tag{6.2}$$

where X is a smooth time-like vector field and b is a smooth real-valued function. Then there exists a pair (A, ϕ) in $L^2_{t,x}((-1, 1) \times B)$ such that the following statements hold:

(1) *There exists a sequence of gauge transforms $\chi^{(n)} \in C_t \mathcal{G}^2((-1, 1) \times B)$ such that, after passing to a subsequence, we have*

$$\left(A^{(n)}_\mu - \partial_\mu \chi^{(n)}, e^{i \chi^{(n)}} \phi^{(n)} \right) \to (A_\mu, \phi) \quad strongly \ in \ L^2_{t,x}((-1, 1) \times B), \tag{6.3}$$

$$\left(F^{(n)}_{\mu\nu}, e^{i \chi^{(n)}} \mathbf{D}^{(n)}_\mu \phi^{(n)} \right) \to (F_{\mu\nu}, \mathbf{D}_\mu \phi) \quad strongly \ in \ L^2_{t,x}((-1, 1) \times B), \tag{6.4}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $\mathbf{D}_\mu \phi = \partial_\mu \phi + i A_\mu \phi$ are defined in the sense of distributions.

(2) *The limiting pair (A, ϕ) is a* weak solution *to* (MKG) *on* $(-1, 1) \times B$, *in the sense of Definition* 6.6 *below. The connection 1-form A obeys, in the sense of distributions, the Coulomb gauge condition*

$$\partial^\ell A_\ell = 0 \quad \text{on } (-1, 1) \times B. \tag{6.5}$$

(3) *The pair (A, ϕ) possesses the following additional regularity:*

$$A \in H^1_{t,x}((-1, 1) \times B), \quad F_{\mu\nu}, \in H^{\frac{1}{2}}_{t,x}((-1, 1) \times B),$$
$$\phi \in H^{\frac{3}{2}}_{t,x}((-1, 1) \times B). \tag{6.6}$$

(4) *Moreover, the pair (A, ϕ) is* stationary *with respect to X, in the sense that*

$$\iota_X F = 0, \quad (\mathbf{D}_X + b)\phi = 0 \quad \text{on } (-1, 1) \times B. \tag{6.7}$$

As a result of taking limits, the notion of weak solutions to (MKG) arises naturally from Proposition 6.1. For our application in Sect. 8, we also need to formulate the notion of locally defined weak solutions $(A_{[\alpha]}, \phi_{[\alpha]})$ that can be pieced together to form a global pair (weak compatible pairs). Developing a theory of these objects is another goal of this section.

Remark 6.2 We remark that weak solutions and their gauge structure play only an auxiliary role in our work. Indeed, the stationarity equation (6.7), combined with (MKG) and the additional regularity (6.6) of (A, ϕ), allow us to infer smoothness of (A, ϕ) via elliptic regularity. This issue is considered in Sect. 7, where we study stationary and self-similar solutions to (MKG).

Remark 6.3 It is in fact possible to obtain stronger convergence than (6.3) namely $A_\mu^{(n)} - \partial_\mu \chi^{(n)} \to A_\mu$ and $e^{i\chi^{(n)}} \phi^{(n)} \to \phi$ in $H^1_{t,x}((-1, 1) \times B)$. Moreover, the limit A_μ obeys the additional regularity $H^{3/2-\varepsilon}_{t,x}((-1, 1) \times B)$ for any $\varepsilon > 0$. As these facts are not necessary for the proof of our main theorem, we omit their proofs to avoid lengthening the paper.

The rest of this section is structured as follows. We first give a proof of Proposition 6.1 in Sect. 6.1, except the statement that the limit (A, ϕ) is a weak solution to (MKG). In Sect. 6.2, we formulate a notion of weak solutions to (MKG) that will be used in our proof. Finally, in Sect. 6.3, we introduce and discuss the notions of smooth and weak compatible pairs, which are local descriptions of smooth and weak solutions to (MKG), respectively.

6.1 Proof of Proposition 6.1

Here we prove Proposition 6.1 modulo the assertion that the limit (A, ϕ) is a weak solution to (MKG), which would be clear once we define the notion of a weak solution in Definition 6.6 below.

Proof The basic idea behind proof is as in [33, Proposition 5.1]: Small energy (6.1) implies local uniform S^1 bound on $(-2, 2) \times 2B$, which can be combined with asymptotic stationarity (6.2) via a microlocal decomposition to conclude strong convergence in $(-1, 1) \times B$. In implementing this strategy, we need to take into account the presence of the constraint equation and the system nature of (MKG) (especially the Maxwell part). Our proof proceeds in several steps.
Step 1 In this step, we use the excision and gluing technique to produce gauge equivalent Coulomb solutions on the smaller region $(-2, 2) \times 2B$, which enjoy a uniform S^1 bound.

Let $(a_j^{(n)}, e_j^{(n)}, f^{(n)}, g^{(n)}) = (A_j^{(n)}, F_{0j}, \phi^{(n)}, \mathbf{D}_t^{(n)} \phi^{(n)}) \restriction_{\{t=0\}}$ be the data for (A, ϕ) on $\{t = 0\}$. Applying Proposition 4.4 to $8B \setminus 4\overline{B}$, we obtain an initial data set $(\widetilde{a}^{(n)}, \widetilde{e}^{(n)}, \widetilde{f}^{(n)}, \widetilde{g}^{(n)}) \in \mathcal{H}^1(\mathbb{R}^4)$ such that $(\widetilde{a}^{(n)}, \widetilde{e}^{(n)}, \widetilde{f}^{(n)}, \widetilde{g}^{(n)}) = (a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)})$ on $4B$ and

$$\mathcal{E}\left[\widetilde{a}^{(n)}, \widetilde{e}^{(n)}, \widetilde{f}^{(n)}, \widetilde{g}^{(n)}\right] \lesssim \epsilon_0^2.$$

by (4.6) and (6.1). Choosing ϵ_0 appropriately, we may ensure that the left-hand side is smaller than ϵ_*^2, which is the threshold for Theorem 4.1.

To pass to the global Coulomb gauge, consider the gauge transformation $\underline{\chi}^{(n)} \in \mathcal{G}^2(\mathbb{R}^4)$ defined by $\underline{\chi}^{(n)} = \triangle^{-1} \partial^\ell \widetilde{a}_\ell^{(n)}$ and let

$$\left(\check{a}^{(n)}, \check{e}^{(n)}, \check{f}^{(n)}, \check{g}^{(n)}\right) := \left(\widetilde{a}^{(n)} - d\underline{\chi}^{(n)}, \widetilde{e}^{(n)}, e^{i\underline{\chi}^{(n)}} \widetilde{f}^{(n)}, e^{i\underline{\chi}^{(n)}} \widetilde{g}^{(n)}\right).$$

This initial data set agrees with $(a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)})$ on $4B$ up to a gauge transformation, i.e.,

$$\left(\check{a}^{(n)}, \check{e}^{(n)}, \check{f}^{(n)}, \check{g}^{(n)}\right) = \left(a^{(n)} - d\underline{\chi}^{(n)}, e^{(n)}, e^{i\underline{\chi}^{(n)}} f^{(n)}, e^{i\underline{\chi}^{(n)}} g^{(n)}\right) \quad \text{on } 4B,$$
$$(6.8)$$

and furthermore obeys the small energy condition

$$\mathcal{E}\left[\check{a}^{(n)}, \check{e}^{(n)}, \check{f}^{(n)}, \check{g}^{(n)}\right] < \epsilon_*^2. \tag{6.9}$$

By small energy global well-posedness (Theorem 4.1), it follows that there exists a unique $C_t \mathcal{H}^1$ admissible solution $(\check{A}^{(n)}, \check{\phi}^{(n)})$ on $\mathbb{R}^{1+4}$ with initial data $(\check{a}^{(n)}, \check{e}^{(n)}, \check{f}^{(n)}, \check{g}^{(n)})$, which obeys

$$\|\check{A}_0^{(n)}\|_{Y^1(\mathbb{R}^{1+4})} + \|\check{A}_x^{(n)}\|_{S^1(\mathbb{R}^{1+4})} + \|\check{\phi}^{(n)}\|_{S^1(\mathbb{R}^{1+4})} \lesssim \epsilon_*. \qquad (6.10)$$

Moreover, by geometric uniqueness (Proposition 4.6) and the simple fact that

$$(-2, 2) \times 2B \subseteq \mathcal{D}^+(\{0\} \times 4B) \cup \mathcal{D}^-(\{0\} \times 4B),$$

there exists $\chi^{(n)} \in C_t \mathcal{G}^2((-2, 2) \times 2B)$ such that

$$\left(\check{A}^{(n)}, \check{\phi}^{(n)}\right) = \left(A^{(n)} - d\chi^{(n)}, e^{i\chi^{(n)}}\phi^{(n)}\right) \quad \text{on } (-2, 2) \times 2B. \quad (6.11)$$

Let $\eta_0, \ldots, \eta_3 \in C_0^\infty(\mathbb{R}^{1+4})$ be such that

$$\eta_j = 1 \text{ on } (-1, 1) \times B, \quad \operatorname{supp} \eta_j \subseteq (-2, 2) \times 2B, \quad \eta_j \eta_{j+1} = \eta_j.$$

for $j = 0, 1, 2, 3$ (except for the last property, for which $j = 0, 1, 2$), which will be fixed for the rest of the proof. We will also often write $\eta = \eta_0$ and $\widetilde{\eta} = \eta_3$. By (6.10) and Remark 4.2, the solution $(\check{A}^{(n)}, \check{\phi}^{(n)})$ satisfies

$$\|\partial_{t,x}\left(\eta_j \check{A}^{(n)}\right)\|_{L_t^\infty L_x^2} + \|\partial_{t,x}\left(\eta_j \check{\phi}^{(n)}\right)\|_{L_t^\infty L_x^2} \lesssim_{\eta_j} \epsilon_0, \qquad (6.12)$$

$$\|\partial_{t,x}\left(\eta_j \check{A}_0^{(n)}\right)\|_{L_t^2 \dot{H}_x^{\frac{1}{2}}} + \|\Box\left(\eta_j \check{A}_x^{(n)}\right)\|_{L_t^2 \dot{H}_x^{-\frac{1}{2}}}$$

$$+ \|\Box\left(\eta_j \check{\phi}^{(n)}\right)\|_{L_t^2 \dot{H}_x^{-\frac{1}{2}}} \lesssim_{\eta_j} \epsilon_0. \qquad (6.13)$$

for any $j = 0, 1, 2, 3$. In particular, in view of (6.12) and Hölder's inequality, the sequence $(\widetilde{\eta}\check{A}^{(n)}, \widetilde{\eta}\check{\phi}^{(n)})$ is uniformly bounded in $H_{t,x}^1$. By the Rellich-Kondrachov theorem, there exists a subsequence, which we still denote by $(\widetilde{\eta}\check{A}^{(n)}, \widetilde{\eta}\check{\phi}^{(n)})$, and a pair $(A, \phi) \in H_{t,x}^1$ such that

$$\left(\widetilde{\eta}\check{A}^{(n)}, \widetilde{\eta}\check{\phi}^{(n)}\right) \rightharpoonup (A, \phi) \quad \text{in } H_{t,x}^1, \quad \left(\widetilde{\eta}\check{A}^{(n)}, \widetilde{\eta}\check{\phi}^{(n)}\right) \to (A, \phi) \quad \text{in } L_{t,x}^2,$$
$$\qquad (6.14)$$

as $n \to \infty$, where the notation $\rightharpoonup$ refers to weak convergence.

Step 2 In this preparatory step, we make a microlocal decomposition of η that will allows us to combine (6.2) with the bound (6.13) on the sequence; see (6.15).

We use the classical pseudo-differential calculus. Let $q_0(\tau, \xi) \in S^0$ be a smooth cutoff such that $q_0 = 1$ to the region $\{(\tau, \xi) : |\tau| \leq (1 - \delta)|\xi|\}$ in Fourier space and $\operatorname{supp} q_0 \subseteq \{(\tau, \xi) : |\tau| \leq (1 - \delta/2)|\xi|\}$, where $\delta > 0$ is to be chosen shortly. On the support of q_0, the norm on the left-hand side of (6.13)

is effective. On the other hand, since $X = X^\mu \partial_\mu$ is a time-like vector field, we have $|X^0(t, x)|^2 > \sum_{j=1}^4 |X^j(t, x)|^2$ everywhere. As $\operatorname{supp} \eta$ is compact, we may choose $\delta > 0$ sufficiently small so that

$$(1 - \delta)^2 |X^0(t, x)| \geq \left(\sum_{j=1}^4 |X^j(t, x)|^2 \right)^{\frac{1}{2}} \quad \text{for } (t, x) \in \operatorname{supp} \eta.$$

With such a choice of $\delta > 0$, the symbol $X^0(t, x)\tau + X^\ell(t, x)\xi_\ell \in S^1$ is *elliptic* on the phase space support of $\eta(t, x)(1 - q_0)(\tau, \xi)$, in the sense that

$$|X^0(t, x)\tau + X^\ell(t, x)\xi_\ell| \geq |X^0(t, x)\tau| - |X^\ell(t, x)\xi_\ell| \geq c_{\delta, \eta, X^0}(|\tau| + |\xi|)$$

for $(t, x) \in \operatorname{supp} \eta$ and $(\tau, \xi) \in \operatorname{supp}(1 - q_0)$, where we may take

$$c_{\delta, \eta, X^0} = \frac{\delta(1 - \delta)}{2} \inf_{\operatorname{supp} \eta} |X^0| > 0.$$

Using the standard construction of a pseudo-differential elliptic parametrix, we may write

$$\eta(1 - q_0)(D_{t,x}) = q_{-1}(t, x, D_{t,x}) \eta X^\mu \partial_\mu + \widetilde{r}_{-1}(t, x, D_{t,x})$$

where $q_{-1}, \widetilde{r}_{-1} \in S^{-1}$. Rearranging the terms, commuting $\eta(t, x)$ with q_0 and applying multiplication by η_1 on the right, we arrive at the decomposition

$$\eta = q_{-1}(t, x, D_{t,x}) \eta X^\mu \partial_\mu + q_0 \eta + r_{-1}(t, x, D_{t,x}) \eta_1, \qquad (6.15)$$

where $r_{-1} \in S^{-1}$ is the sum of $\widetilde{r}_{-1}$ and the commutator between η and q_0.

Step 3 Here we show the strong convergence $\eta F_{\mu\nu}^{(n)} \to \eta F_{\mu\nu}$ in $L_{t,x}^2$, where we remind the reader that $F_{\mu\nu} = \hat{F}_{\mu\nu}$ by gauge invariance of the curvature 2-form. By (6.15), we may write

$$\eta F_{\mu\nu}^{(n)} = q_{-1}(t, x, D_{t,x}) \eta X^\lambda \partial_\lambda F_{\mu\nu}^{(n)} + q_0(D_{t,x}) \eta F_{\mu\nu}^{(n)} + r_{-1}(t, x, D_{t,x}) \eta_1 F_{\mu\nu}^{(n)}.$$

Using $dF^{(n)} = 0$, we rewrite $\eta X^\lambda \partial_\lambda F_{\mu\nu}^{(n)}$ as

$$\eta X^\lambda \partial_\lambda F_{\mu\nu}^{(n)} = \partial_\mu \left(\eta X^\lambda F_{\lambda\nu}^{(n)} \right) - \partial_\nu (\eta X^\lambda F_{\lambda\mu}^{(n)})$$
$$- \partial_\mu (\eta X^\lambda) F_{\lambda\nu}^{(n)} + \partial_\nu \left(\eta X^\lambda \right) F_{\lambda\mu}^{(n)},$$

and hence we arrive at

$$\eta F_{\mu\nu}^{(n)} = q_{-1}(t, x, D_{t,x}) \left[\partial_\mu (\eta (\iota_X F^{(n)})_\nu) - \partial_\nu (\eta (\iota_X F^{(n)})_\mu) \right] + R_M[F^{(n)}]_{\mu\nu}$$
$$(6.16)$$

where

$$R_{\mathrm{M}}[F^{(n)}]_{\mu\nu} = q_0(D_{t,x})\eta F^{(n)}_{\mu\nu} - q_{-1}(t, x, D_{t,x})\big[\partial_\mu(\eta X^\lambda)F^{(n)}_{\lambda\nu}$$
$$-\partial_\nu(\eta X^\lambda)F^{(n)}_{\lambda\mu}\big] + r_{-1}(t, x, D_{t,x})\eta_1 F^{(n)}_{\mu\nu}.$$

By (6.2), it follows that

$$\|q_{-1}(t, x, D_{t,x})\big[\partial_\mu\left(\eta(\iota_X F^{(n)})_\nu\right) - \partial_\nu\left(\eta(\iota_X F^{(n)})_\mu\right)\big]\|_{L^2_{t,x}} \to 0.$$

Moreover, we claim that $R_{\mathrm{M}}[F^{(n)}]_{\mu\nu}$ enjoys improved regularity, i.e.,

$$\|R_{\mathrm{M}}[F^{(n)}]_{\mu\nu}\|_{H^{\frac{1}{2}}_{t,x}} \lesssim \epsilon_0 \quad \text{uniformly in } n. \tag{6.17}$$

By the Rellich-Kondrachov theorem, after passing to a subsequence of $(\check{A}^{(n)}, \check{\phi}^{(n)})$, the sequence $\widetilde{\eta}R_{\mathrm{M}}[F^{(n)}]_{\mu\nu}$ is strongly convergent in $L^2_{t,x}$; moreover, we can also ensure that the limit belongs to $H^{\frac{1}{2}}_{t,x}$. Combining these facts, as well as the identity $\eta\widetilde{\eta} = \eta$, we see that $\eta F^{(n)}_{\mu\nu}$ is strongly convergent in $L^2_{t,x}$ to a limit that belongs to $H^{\frac{1}{2}}_{t,x}$. Since $\widetilde{\eta}\check{A}_\mu \to A_\mu$ in $L^2_{t,x}$, the limit is equal to $\eta F_{\mu\nu}$. Hence the statements regarding F in (6.4) and (6.6) follow.

It remains to verify the claim (6.17); it is at this point we use the uniform bounds (6.12) and (6.13). Using the formula $\eta F^{(n)} = \eta(\mathrm{d}\check{A}^{(n)}) = \mathrm{d}(\eta\check{A}^{(n)}) - \mathrm{d}\eta \wedge \check{A}^{(n)}$ and the support property of the symbol q_0, we obtain

$$\|q_0(D_{t,x})\eta F^{(n)}\|_{H^{\frac{1}{2}}_{t,x}} \lesssim \|q_0(D_{t,x})\mathrm{d}\left(\eta\check{A}^{(n)}\right)\|_{H^{\frac{1}{2}}_{t,x}}$$
$$+ \|q_0(D_{t,x})\left(\mathrm{d}\eta \wedge \check{A}^{(n)}\right)\|_{H^{\frac{1}{2}}_{t,x}}.$$

The second term on the right-hand side is bounded by ϵ_0 thanks to (6.12). To handle the first term, we divide the space-time Fourier space into the regions $\{|\tau| + |\xi| \le 1\}$ and $\{|\tau| + |\xi| > 1\}$. Also distinguishing the temporal and spatial components of $\check{A}^{(n)}$, we may estimate

$$\|q_0(D_{t,x})\mathrm{d}\left(\eta\check{A}^{(n)}\right)\|_{H^{\frac{1}{2}}_{t,x}} \lesssim \|\partial_{t,x}\left(\eta\check{A}^{(n)}_0\right)\|_{L^2_t \dot{H}^{1/2}_x}$$
$$+ \|\Box(\eta\check{A}^{(n)}_x)\|_{L^2_t \dot{H}^{-1/2}_x} + \|\eta\check{A}^{(n)}\|_{L^2_{t,x}}$$

Using (6.13) for the first two terms and (6.12) for the last, the entire right-hand side is bounded by ϵ_0.

For the remainder $R_M[F^{(n)}]_{\mu\nu} - q_0(D_{t,x})\eta F_{\mu\nu}^{(n)}$, we begin by observing that $\|\eta_2 F^{(n)}\|_{L_{t,x}^2} \lesssim \epsilon_0$ by the formula $F^{(n)} = d\check{A}^{(n)}$ and (6.12). Then we have

$$\|R_M[F^{(n)}]_{\mu\nu} - q_0(D_{t,x})\eta F_{\mu\nu}^{(n)}\|_{H_{t,x}^1} \lesssim \|\eta_2 F^{(n)}\|_{L_{t,x}^2} \lesssim \epsilon_0,$$

which proves the claim (6.17).

Step 4 In this intermediate step, we use strong $L_{t,x}^2$ convergence of $F_{\mu\nu}^{(n)}$ to prove

$$\eta \check{A}_\mu^{(n)} \to \eta A_\mu \quad \text{strongly in } L_t^2 H_x^1, \tag{6.18}$$

as $n \to \infty$, up to a subsequence. We also prove improved regularity for the limit A_μ, i.e.,

$$\partial_x(\eta A_\mu) \in H_{t,x}^{\frac{1}{2}}. \tag{6.19}$$

To begin with, observe that $\triangle \check{A}_\mu^{(n)} = \partial^\ell F_{\ell\mu}^{(n)}$ by the Coulomb gauge condition. Therefore, for each spatial component $\mu = k \in \{1, 2, 3, 4\}$, we have

$$\eta \check{A}_k^{(n)} = \triangle^{-1}\left(\partial^\ell\left(\eta F_{\ell k}^{(n)}\right) + [\triangle, \eta]\check{A}_k^{(n)} + \left[\eta, \partial^\ell\right]F_{\ell k}^{(n)}\right). \tag{6.20}$$

For any $j \in \{1, 2, 3, 4\}$, note that $\partial_j \triangle^{-1}\partial^\ell(\eta F_{\ell k}^{(n)})$ is strongly convergent in $L_{t,x}^2$, thanks to the previous step. Writing out $F^{(n)} = d\check{A}^{(n)}$ and using the strong $L_{t,x}^2$ convergence of $\tilde{\eta}\check{A}_k^{(n)}$, it follows that the remainder $\partial_j \triangle^{-1}([\triangle, \eta]\check{A}_k^{(n)} + [\eta, \partial^\ell]F_{\ell k}^{(n)})$ is strongly convergent in $L_{t,x}^2$ as well. Hence (6.18) holds for $\mu \in \{1, 2, 3, 4\}$.

In the case $\mu = 0$, note that (6.12) and (6.13) already imply

$$\left\|\partial_x\left(\tilde{\eta}\check{A}_0^{(n)}\right)\right\|_{H_{t,x}^{\frac{1}{2}}} \lesssim \epsilon_0 \quad \text{uniformly in } n. \tag{6.21}$$

Therefore, after taking a suitable subsequence, the desired convergence (6.18) (by the Rellich-Kondrachov theorem) as well as the improved regularity (6.19) follow.

It only remains to prove the improved regularity (6.19) for $\mu = k \in \{1, 2, 3, 4\}$. First, by (6.20) and the improved regularity $\eta F \in H_{t,x}^{\frac{1}{2}}$, $\tilde{\eta}\check{A} \in H_{t,x}^1$, it follows that $\eta \check{A}_k \in L_t^2 H_x^{\frac{3}{2}}$. Then using the identity

$$\partial_t(\eta A_k) - \partial_k(\eta A_0) = \eta F_{0k} + [\partial_j, \eta]A_k - [\partial_k, \eta]A_0,$$

and the improved regularity $\partial_x(\eta A_0) \in H^{\frac{1}{2}}_{t,x}$, as well as $\eta F \in H^{\frac{1}{2}}_{t,x}$, $\widetilde{\eta}\check{A} \in H^1_{t,x}$, we have $\partial_t(\eta\check{A}_k) \in H^{\frac{1}{2}}_{t,x}$. It follows that $\eta\check{A}_k \in H^{\frac{3}{2}}_{t,x}$, which is better than what we need.

Step 5 In this step, we show that $\eta\check{\mathbf{D}}^{(n)}\check{\phi}^{(n)} \to \eta\mathbf{D}\phi$ in $L^2_{t,x}$ and $\eta\phi \in H^{\frac{3}{2}}_{t,x}$. For the former, from the decomposition

$$\eta\check{\mathbf{D}}^{(n)}_\mu\check{\phi}^{(n)} = \eta\partial_\mu\check{\phi}^{(n)} + i\eta\check{A}^{(n)}_\mu\check{\phi}^{(n)},$$

the convergence $\eta\check{A}^{(n)}_\mu \to \eta A$ in $L^2_t H^1_x$ and (6.12), we see that it suffices to prove

$$\eta\partial_\mu\check{\phi}^{(n)} \to \eta\partial_\mu\phi \quad \text{in } L^2_{t,x}. \tag{6.22}$$

By (6.15), we have

$$\eta\check{\phi}^{(n)} = q_{-1}(t, x, D_{t,x})\eta X^\mu\partial_\mu\check{\phi}^{(n)} + q_0\eta\check{\phi}^{(n)} + r_{-1}(t, x, D_{t,x})\eta_1\check{\phi}^{(n)}$$

To use (6.2), we rewrite $\eta X^\mu\partial_\mu\check{\phi}^{(n)}$ as

$$\eta X^\mu\partial_\mu\check{\phi}^{(n)} = \eta(\check{\mathbf{D}}^{(n)}_X + b)\check{\phi}^{(n)} - iX^\nu\check{A}^{(n)}_\nu\eta\check{\phi}^{(n)} - \eta b\check{\phi}^{(n)}.$$

where $\check{\mathbf{D}}^{(n)} = \mathrm{d} + i\check{A}^{(n)}$. Expanding $\eta\check{A}^{(n)} = \eta(\check{A}^{(n)} - A^{(n)}) + \eta A^{(n)}$, we arrive at

$$\begin{aligned}
\eta\check{\phi}^{(n)} = {}& q_{-1}(t, x, D_{t,x})\eta\left(\check{\mathbf{D}}^{(n)}_X + b\right)\check{\phi}^{(n)} \\
& - iq_{-1}(t, x, D_{t,x})X^\nu\eta\left(\check{A}^{(n)}_\nu - A_\nu\right)\check{\phi}^{(n)} \\
& - iq_{-1}(t, x, D_{t,x})X^\nu\eta A_\nu\check{\phi}^{(n)} + R_{KG}\left[\check{\phi}^{(n)}\right]
\end{aligned} \tag{6.23}$$

where

$$R_{KG}[\check{\phi}^{(n)}] := q_0\eta\check{\phi}^{(n)} + r_{-1}(t, x, D_{t,x})\eta_1\check{\phi}^{(n)} - q_{-1}(t, x, D_{t,x})b\eta\check{\phi}^{(n)}.$$

As in Step 2, for the first term we have

$$\|q_{-1}(t, x, D_{t,x})\eta(\check{\mathbf{D}}^{(n)}_X + b)\check{\phi}^{(n)}\|_{H^1_{t,x}} \to 0$$

as $n \to \infty$, thanks to (6.2). For the second term, we have

$$\|q_{-1}(t, x, D_{t,x})X^\nu \eta(\check{A}_\nu^{(n)} - A_\nu)\check{\phi}^{(n)}\|_{H^1_{t,x}}$$
$$\lesssim \|\eta(\check{A}_\nu^{(n)} - A_\nu)\|_{L^2_t L^4_x}\|\check{\phi}^{(n)}\|_{L^\infty_t L^4_x} \to 0$$

as $n \to \infty$, by Hölder, Sobolev in x, $L^2_t H^1_x$ convergence of $\eta \check{A}_\nu^{(n)}$ to ηA_ν and (6.12). On the other hand, for the third term, we have

$$\|q_{-1}(t, x, D_{t,x})X^\nu \eta A_\nu \check{\phi}^{(n)}\|_{H^{\frac{3}{2}}_{t,x}} \lesssim \epsilon_0 \|\langle D_x \rangle \langle D_{t,x} \rangle^{\frac{1}{2}}(\eta A)\|_{L^2_{t,x}} \text{ uniformly in } n.$$

where we used Lemma 6.4 below with $f = \eta A_\nu$ and $g = \check{\phi}^{(n)}$. We also used the obvious bound $\|\eta A_\nu \check{\phi}^{(n)}\|_{L^2_{t,x}} \lesssim \epsilon_0 \|\langle D_x \rangle (\eta A)\|_{L^2_{t,x}}$, which follows from Hölder, Sobolev in x and (6.12), to control the $L^2_{t,x}$ norm of the left-hand side. Finally, for $R_{\mathrm{KG}}[\check{\phi}^{(n)}]$ we have, as in Step 3,

$$\|R_{\mathrm{KG}}\left[\check{\phi}^{(n)}\right]\|_{H^{\frac{3}{2}}_{t,x}} \lesssim \epsilon_0 \quad \text{uniformly in } n.$$

By the Rellich-Kondrachov theorem, there exists a subsequence (which we still denote by $\check{\phi}^{(n)}$) such that

$$\tilde{\eta}\left(-iq_{-1}(t, x, D_{t,x})X^\nu \eta A_\nu \check{\phi}^{(n)} + R_{\mathrm{KG}}\left[\check{\phi}^{(n)}\right]\right)$$

is strongly convergent in $H^1_{t,x}$ to a limit that belongs to $H^{\frac{3}{2}}_{t,x}$. As a consequence of these facts, as well as the identity $\eta\tilde{\eta} = \eta$, it follows that $\eta\check{\phi}^{(n)}$ is strongly convergent in $H^1_{t,x}$ to a limit in $H^{\frac{3}{2}}_{t,x}$. Finally, since $\tilde{\eta}\check{\phi}^{(n)} \to \phi$ in $L^2_{t,x}$, the limit is equal to $\eta\phi$. $\qquad\square$

Lemma 6.4 *For $f, g \in \mathcal{S}(\mathbb{R}^{1+4})$, we have*

$$\|fg\|_{\dot{H}^{\frac{1}{2}}_{t,x}} \lesssim \||D_{t,x}|^{\frac{1}{2}}f\|_{L^2_t \dot{H}^1_x}\|D_{t,x}g\|_{L^\infty_t L^2_x}. \tag{6.24}$$

Proof We use the Littlewood-Paley projections $\{S_j\}$ in $\mathbb{R}^{1+4}$. For every $j \in \mathbb{Z}$, we decompose

$$S_j(fg) = S_j\left((S_{>j-10}f)g\right) + S_j\left(S_{\leq j-10}f S_{[j-5,j+5]}g\right)$$

Using Sobolev and Hölder, we estimate each term on the right-hand side as follows:

$$\|S_j((S_{>j-10}f)g)\|_{\dot{H}^{\frac{1}{2}}_{t,x}}$$

$$\lesssim \sum_{j_1>j-10} 2^{\frac{1}{2}j}\|S_{j_1}f\|_{L^2_t L^4_x}\|g\|_{L^\infty_t L^4_x}$$

$$\lesssim \|D_{t,x}g\|_{L^\infty_t L^2_x} \sum_{j_1>j-10} 2^{\frac{1}{2}(j-j_1)}\||D_{t,x}|^{\frac{1}{2}}S_{j_1}f\|_{L^2_t \dot{H}^1_x},$$

$$\|S_j(S_{\leq j-10}f\, S_{[j-5,j+5]}g)\|_{\dot{H}^{\frac{1}{2}}_{t,x}}$$

$$\lesssim \sum_{j_1\leq j-10} 2^{\frac{1}{2}j}\|S_{j_1}f\|_{L^2_t L^\infty_x}\|S_{[j-5,j+5]}g\|_{L^\infty_t L^2_x}$$

$$\lesssim \|D_{t,x}g\|_{L^\infty_t L^2_x} \sum_{j_1\leq j-10} 2^{\frac{1}{2}(j_1-j)}\||D_{t,x}|^{\frac{1}{2}}S_{j_1}f\|_{L^2_t \dot{H}^1_x}.$$

Thanks to the exponential gain $2^{-\frac{1}{2}|j-j_1|}$, we have

$$\sum_j \|S_j(fg)\|^2_{\dot{H}^{\frac{1}{2}}_{t,x}} \lesssim \|D_{t,x}g\|^2_{L^\infty_t L^2_x} \sum_{j_1} \||D_{t,x}|^{\frac{1}{2}}S_{j_1}f\|^2_{L^2_t \dot{H}^1_x}.$$

The desired estimate is now a consequence of almost orthogonality of $\{S_j\}_{j\in\mathbb{Z}}$ in $L^2_{t,x}$.   □

6.2 Weak solutions to (MKG)

We first define a function space that is suitable for a weak formulation of (MKG).

Definition 6.5 Let $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ be an open set. We define $\mathcal{X}^w(\mathcal{O})$ to be the linear space of pairs (A, ϕ), where A is a real-valued 1-form and ϕ is a $\mathbb{C}$-valued function on $\mathcal{O}$, such that

$$A_\mu, \phi \in L^2_{t,x}(\mathcal{O}), \ F_{\mu\nu}, \mathbf{D}_\mu\phi \in L^2_{t,x}(\mathcal{O}) \quad \text{for all } \mu, \nu = 0, 1, \ldots, 4, \quad (6.25)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $\mathbf{D}_\mu\phi = \partial_\mu\phi + iA_\mu\phi$ in the sense of distributions.

We may now define a notion of weak solutions to (MKG) as follows.

Definition 6.6 *(Weak solutions to* (MKG)*)* Let $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ be an open set, and let $(A, \phi) \in \mathcal{X}^w(\mathcal{O})$. We say that (A, ϕ) is a weak solution to (MKG) on $\mathcal{O}$ if for every real-valued 1-form $\omega \in C^\infty_0(\mathcal{O})$ and complex-valued function $\varphi \in C^\infty_0(\mathcal{O})$, we have

$$\iint_{\mathcal{O}} F_{\nu\mu} \partial^\mu \omega^\nu + \mathrm{Im}(\phi \overline{\mathbf{D}_\nu \phi}) \omega^\nu \, \mathrm{d}t \mathrm{d}x = 0, \tag{6.26}$$

$$\iint_{\mathcal{O}} \mathrm{Re}(\mathbf{D}_\mu \phi \overline{\partial^\mu \varphi}) + \mathrm{Im}(A^\mu \mathbf{D}_\mu \phi \overline{\varphi}) \, \mathrm{d}t \mathrm{d}x = 0. \tag{6.27}$$

By an integration by parts argument, it may be readily verified that admissible and classical solutions to (MKG) are indeed weak solutions. In the converse direction, if (A, ϕ) is a weak solution to (MKG) that is furthermore smooth, then (A, ϕ) solves (MKG) in the usual, classical sense.

Next, we discuss the gauge structure of weak solutions to (MKG). We first define the space of gauge transformations between pairs in $\mathcal{X}^w$.

Definition 6.7 Given an open set $\mathcal{O} \subseteq \mathbb{R}^{1+4}$, let $\mathcal{Y}^w(\mathcal{O})$ be the space of real-valued functions χ on $\mathcal{O}$ such that $\chi \in H^1_{t,x}(\mathcal{O})$.

Indeed, note that if $(A, \phi) \in \mathcal{X}^w$ and $\chi \in \mathcal{Y}^w$, then the gauge transform $(\widetilde{A}, \widetilde{\phi}) := (A - \mathrm{d}\chi, e^{i\chi})$ also belongs to $\mathcal{X}^w$. Moreover, if (A, ϕ) is a weak solution to (MKG) then so is $(\widetilde{A}, \widetilde{\phi})$, as the next lemma demonstrates.

Lemma 6.8 *Let $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ be an open set, and let $(A, \phi) \in \mathcal{X}^w(\mathcal{O})$ be a weak solution to (MKG). Then for every $\chi \in \mathcal{Y}^w(\mathcal{O})$, the gauge transform $(\widetilde{A}, \widetilde{\phi}) := (A - \mathrm{d}\chi, e^{i\chi}\phi)$ also belongs to $\mathcal{X}^w(\mathcal{O})$ and is a weak solution to (MKG).*

Proof We need to verify (6.26) and (6.27) for $(\widetilde{A}, \widetilde{\phi})$. For (6.26) there is nothing to verify, as both F and $\mathrm{Im}(\phi \overline{\mathbf{D}\phi})$ are invariant under gauge transformation. For (6.27), we have

$$\iint_{\mathcal{O}} \mathrm{Re}\left(\widetilde{\mathbf{D}}_\mu \widetilde{\phi} \, \overline{\partial^\mu \varphi}\right) + \mathrm{Im}\left(\widetilde{A}^\mu \widetilde{\mathbf{D}}_\mu \widetilde{\phi} \, \overline{\varphi}\right) \, \mathrm{d}t \mathrm{d}x$$
$$= \iint_{\mathcal{O}} \mathrm{Re}\left(\mathbf{D}_\mu \phi \overline{\partial^\mu (e^{-i\chi}\varphi)}\right) + \mathrm{Im}\left(A^\mu \mathbf{D}_\mu \phi \, \overline{e^{-i\chi}\varphi}\right) \, \mathrm{d}t \mathrm{d}x.$$

Observe that if $\chi \in C^\infty(\mathcal{O})$, then the last line would be equal to zero by (6.27) for (A, ϕ). Considering a sequence $\chi^{(n)} \in C^\infty(\mathcal{O})$ such that $\chi^{(n)} \to \chi$ in the $H^1_{t,x}(\mathcal{O})$ topology and also pointwise almost everywhere, it can be seen that the last line is indeed zero, by the dominated convergence theorem, Leibniz's rule and Hölder's inequality. $\square$

6.3 Local description of solutions to (MKG)

Here we discuss how to describe a solution to (MKG) by local data. More precisely, given an open cover $\mathcal{Q} = \{Q_\alpha\}$ of an open set $\mathcal{O} \subseteq \mathbb{R}^{1+4}$, we would like to describe a solution to (MKG) on $\mathcal{O}$ by local solutions on Q_α satisfying

certain compatibility conditions, which ensure that the local solutions combine to form a single solution on $\mathcal{O}$. This idea is made precise by the ensuing definition.

Definition 6.9 *(Smooth compatible pairs)* Let $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ be an open set and let $\mathcal{Q} = \{Q_\alpha\}$ be a locally finite open covering of $\mathcal{O}$. For each index α, consider a pair $(A_{[\alpha]}, \phi_{[\alpha]}) \in C_{t,x}^\infty(Q_\alpha)$, where $A_{[\alpha]}$ is a real-valued 1-form and $\phi_{[\alpha]}$ is a $\mathbb{C}$-valued function on Q_α. We say that $(A_{[\alpha]}, \phi_{[\alpha]})$ are smooth compatible pairs if for every α, β, there exists a gauge transformation $\chi_{[\alpha\beta]} \in C_{t,x}^\infty(Q_\alpha \cap Q_\beta)$ such that the following properties hold:

(1) For every α, we have $\chi_{[\alpha\alpha]} = 0$.
(2) For every α, β, we have

$$(A_{[\beta]}, \phi_{[\beta]}) = (A_{[\alpha]} - \mathrm{d}\chi_{[\alpha\beta]}, e^{i\chi_{[\alpha\beta]}}\phi_{[\alpha]}) \quad \text{on } Q_\alpha \cap Q_\beta. \quad (6.28)$$

(3) For every α, β, γ, the following cocycle condition is satisfied:

$$\chi_{[\alpha\beta]} + \chi_{[\beta\gamma]} + \chi_{[\gamma\alpha]} \in 2\pi\mathbb{Z} \quad \text{on } Q_\alpha \cap Q_\beta \cap Q_\gamma. \quad (6.29)$$

The notion of (gauge-)equivalence of compatible pairs is defined as follows.

Definition 6.10 *(Equivalence of smooth compatible pairs)* Let $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ be an open set, and let $\mathcal{Q} = \{Q_\alpha\}$, $\mathcal{Q}' = \{Q'_\beta\}$ be locally finite open coverings of $\mathcal{O}$. Consider two sets of smooth compatible pairs $(A_{[\alpha]}, \phi_{[\alpha]})$ and $(A'_{[\beta]}, \phi'_{[\beta]})$ on $\mathcal{Q}$ and $\mathcal{Q}'$, respectively. When $\mathcal{Q}'$ is a refinement of $\mathcal{Q}$ (i.e., for every β there exists $\alpha(\beta)$ such that $Q'_\beta \subseteq Q_\alpha$), we say that $(A_{[\alpha]}, \phi_{[\alpha]})$ and $(A'_{[\beta]}, \phi'_{[\beta]})$ are (gauge-)equivalent if for every β there exists $\chi_{[\beta]} \in C_{t,x}^\infty(Q'_\beta)$ such that $(A'_{[\beta]}, \phi'_{[\beta]}) = (A_{[\alpha]} - \mathrm{d}\chi_{[\beta]}, \phi_{[\alpha]}e^{i\chi_{[\beta]}})$. In the general case, we say that $(A_{[\alpha]}, \phi_{[\alpha]})$ and $(A'_{[\beta]}, \phi'_{[\beta]})$ are (gauge-)equivalent if there exists a common refinement $\mathcal{Q}''$ of $\mathcal{Q}, \mathcal{Q}'$ and a set of smooth compatible pairs $(A''_{[\gamma]}, \phi''_{[\gamma]})$ on $\mathcal{Q}''$ which is equivalent to both $(A_{[\alpha]}, \phi_{[\alpha]})$ and $(A'_{[\beta]}, \phi'_{[\beta]})$.

Remark 6.11 In more geometric terms, compatible pairs $(A_{[\alpha]}, \phi_{[\alpha]})$ on Q_α are precisely expressions of a connection A and a section ϕ of a complex line bundle L in local trivializations $L\restriction_{Q_\alpha} \simeq Q_\alpha \times \mathbb{C}$. Moreover, equivalent sets of compatible pairs are alternative expressions of the same global pair (A, ϕ).

In fact, expression of connections and sections in local trivializations in the fashion of Definition 6.9 is necessary if the complex line bundle L under consideration is topologically nontrivial (i.e., L is not homeomorphic to the product of $\mathbb{C}$ and the base space). In our setting, however, there is no loss of generality in simply identifying connections and sections of L with real-valued 1-forms and complex-valued functions, respectively, as all base spaces

we consider (e.g., $\mathcal{O} = I \times \mathbb{R}^4$ or $C_{[T,\infty)}^T$ for some $T > 0$) are contractible and hence all complex line bundles over such spaces are topologically trivial. In this case, every smooth compatible pairs on $\mathcal{O}$ is equivalent to a global smooth pair (A, ϕ) on $\mathcal{O}$.

Remark 6.12 We emphasize that no delicate patching is needed for smooth compatible pairs in this paper, since all we need is merely the soft fact that the energy argument in Sect. 5 and the stress tensor argument in Sect. 7 (which are both gauge invariant) can be justified. In contrast, in [26] an elaborate patching argument had to be developed in order to control the S^1 norm of the equivalent global pair in the Coulomb gauge.

Based on the spaces introduced for the weak formulation of (MKG) discussed above, we can also formulate the notion of weak compatible pairs.

Definition 6.13 *(Weak compatible pairs)* Let $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ be an open set and let $\mathcal{Q} = \{Q_\alpha\}$ be a locally finite covering of $\mathcal{O}$. For each index α, consider a pair $(A_{[\alpha]}, \phi_{[\alpha]}) \in \mathcal{X}^w(Q_\alpha)$. We say that $(A_{[\alpha]}, \phi_{[\alpha]})$ are weak compatible pairs if for every α, β, there exists a gauge transformation $\chi_{[\alpha\beta]} \in \mathcal{Y}^w(Q_\alpha \cap Q_\beta)$ such that the properties (1)–(3) in Definition 6.9 hold almost everywhere.

The notion of equivalent sets of weak compatible pairs is defined as in Definition 6.10, where the space $C_{t,x}^\infty(Q'_\beta)$ is replaced by $\mathcal{Y}^w(Q'_\beta)$.

Geometrically, weak compatible pairs $(A_{[\alpha]}, \phi_{[\alpha]})$ may be thought of as local descriptions of a connection and a section defined on a rough complex line bundle L. A simple but crucial observation is that smoothness of the pairs $(A_{[\alpha]}, \phi_{[\alpha]})$ implies smoothness of the gauge transformations $\chi_{[\alpha\beta]}$. Indeed, simply note that $d\chi_{[\alpha\beta]} = A_{[\alpha]} - A_{[\beta]}$ by the property (2) in Definition 6.9. As this fact will play an important role in our argument (see Proposition 7.3), we record it as a separate lemma.

Lemma 6.14 *Let $\mathcal{Q} = \{Q_\alpha\}$ be an open cover of $\mathcal{O} \subseteq \mathbb{R}^{1+4}$, and let $(A_{[\alpha]}, \phi_{[\alpha]})$ on Q_α be weak compatible pairs. If $A_{[\alpha]}, \phi_{[\alpha]} \in C^\infty(Q_\alpha)$ for every α, then $(A_{[\alpha]}, \phi_{[\alpha]})$ form smooth compatible pairs in the sense of Definition 6.9.*

We end this subsection with another simple lemma, which will be used later to show that the local solutions obtained from Proposition 6.1 in the limit form weak compatible pairs.

Lemma 6.15 *Let $Q_1, Q_2 \subseteq \mathbb{R}^{1+4}$ be open sets such that $Q_1 \cap Q_2 \neq \emptyset$ is an open bounded set with a piecewise smooth boundary. Consider sequences $(A_{[\alpha]}^{(n)}, \phi_{[\alpha]}^{(n)}) \in \mathcal{X}^w(Q_\alpha)$ ($\alpha = 1, 2$) and $\chi_{[12]}^{(n)} \in \mathcal{Y}^w(Q_1 \cap Q_2)$ such that*

$$\left(A_{[2]}^{(n)}, \phi_{[2]}^{(n)} \right) = \left(A_{[1]}^{(n)} - d\chi_{[12]}^{(n)}, \phi_{[1]}^{(n)} e^{i\chi_{[12]}^{(n)}} \right) \quad \text{a.e. on } Q_1 \cap Q_2. \quad (6.30)$$

In other words, $(A^{(n)}_{[\alpha]}, \phi^{(n)}_{[\alpha]})$ are weak compatible pairs for each n. Suppose furthermore that each sequence $(A^{(n)}_{[\alpha]}, \phi^{(n)}_{[\alpha]})$ has a limit $(A_{[\alpha]}, \phi_{[\alpha]})$ in $\mathcal{X}^w(Q_\alpha)$ as $n \to \infty$. Then the limits $(A_{[\alpha]}, \phi_{[\alpha]})$ $(\alpha = 1, 2)$ also form weak compatible pairs, i.e., there exists $\chi_{[12]} \in \mathcal{Y}^w(Q_1 \cap Q_2)$ such that

$$(A_{[2]}, \phi_{[2]}) = (A_{[1]} - d\chi_{[12]}, \phi_{[1]} e^{i\chi_{[12]}}) \quad \text{a.e. on } Q_1 \cap Q_2. \quad (6.31)$$

Moreover, there exists a subsequence of $\chi^{(n)}_{[12]}$ that converges[13] to $\chi_{[12]}$ in $\mathcal{Y}^w(Q_1 \cap Q_2)$ up to integer multiples of 2π.

Proof Let $\overline{\chi}^{(n)}_{[12]} := \int_{Q_1 \cap Q_2} \chi^{(n)}_{[12]}$ denote the mean of $\chi^{(n)}_{[12]}$. By Poincaré's inequality, the identity $d\chi^{(n)}_{[12]} = A^{(n)}_{[1]} - A^{(n)}_{[2]}$ and the $L^2_{t,x}$ convergence of $A^{(n)}_{[\alpha]}(\alpha = 1, 2)$, the mean-zero part $\hat{\chi}^{(n)}_{[12]} := \chi^{(n)}_{[12]} - \overline{\chi}^{(n)}_{[12]}$ converges to a limit $\hat{\chi}_{[12]}$ in $\mathcal{Y}^w(Q_1 \cap Q_2) = H^1_{t,x}(Q_1 \cap Q_2)$. On the other hand, we can easily extract a convergent subsequence from the bounded sequence $e^{i\overline{\chi}^{(n)}_{[12]}}$; abusing the notation a bit, we denote the subsequence still by $e^{i\overline{\chi}^{(n)}_{[12]}}$, and the limit by $e^{i\overline{\chi}_{[12]}}$ for some $\overline{\chi}_{[12]} \in \mathbb{R}$. It follows that $\chi^{(n)}_{[12]}$ converges to $\chi_{[12]} := \hat{\chi}_{[12]} + \overline{\chi}_{[12]}$ in $\mathcal{Y}^w(Q_1 \cap Q_2)$ as $n \to \infty$ up to integer multiples of 2π. The desired gauge equivalence in the limit (6.31) is now an easy consequence of (6.30) and the above convergences.

7 Stationary/self-similar solutions with finite energy

In the context of the blow-up analysis to be performed in Sect. 8, the local strong compactness result (Proposition 6.1) will give rise to two types of solutions to (MKG):

- A stationary solution (A, ϕ), which is defined by the property

$$\iota_Y F = 0, \quad \mathbf{D}_Y \phi = 0 \quad (7.1)$$

  for some constant time-like vector field Y; or
- A self-similar solution (A, ϕ), defined by the property

$$\iota_{X_0} F = 0, \quad \left(\mathbf{D}_{X_0} + \frac{1}{\rho}\right)\phi = 0. \quad (7.2)$$

In Sects. 7.1 and 7.2, we show that such solutions must be trivial under the finite energy assumption. We use the method of stress tensor, which is the

---

[13] That is, there exists $k_m \in \mathbb{Z}$ such that $\chi^{(n_m)}_{[12]} + 2\pi k_m \to \chi_{[12]}$ in $\mathcal{Y}^w(Q_1 \cap Q_2)$.

elliptic version of the energy-momentum-stress tensor considered in Sect. 5. In Sect. 7.3, we establish an elliptic regularity result for these solutions under the improved regularity assumption (6.6) ensured by Proposition 6.1.

7.1 Triviality of finite energy stationary solutions

As any unit constant time-like vector field Y can be Lorentz transformed to the vector field $T = \partial_t$ in the rectilinear coordinates, we may assume that $Y = T$. Our main result in this case is as follows.

Proposition 7.1 *Let (A, ϕ) be a smooth solution to* (MKG) *on $\mathbb{R}^{1+4}$ with $\iota_T F = 0$ and $\mathbf{D}_T \phi = 0$. Suppose furthermore that (A, ϕ) has finite energy, i.e., $\mathcal{E}_{\{0\}\times\mathbb{R}^4}[A, \phi] < \infty$. Then $\mathcal{E}_{\{0\}\times\mathbb{R}^4}[A, \phi] = 0$.*

Proof We use the rectilinear coordinates $(t = x^0, x^1, \ldots, x^4)$, in which $T = \partial_t$. By the stationarity assumptions $(\iota_T F)(\partial_j) = F_{0j} = 0$ and $\mathbf{D}_T \phi = \mathbf{D}_0 \phi = 0$, (MKG) reduces to the following elliptic system on each constant t hypersurface:

$$\begin{cases} \partial^\ell F_{j\ell} = \text{Im}(\phi \overline{\mathbf{D}_j \phi}), \\ \mathbf{D}^\ell \mathbf{D}_\ell \phi = 0. \end{cases} \tag{7.3}$$

Henceforth, we work with F, ϕ restricted to the hypersurface $\{t = 0\}$.

For the purpose of showing $\mathcal{E}[A, \phi] = 0$, consider the following stress tensor associated to (7.3):

$$\mathcal{Q}_{jk}[A, \phi] := \text{Re}(\mathbf{D}_j \phi \overline{\mathbf{D}_k \phi}) - \frac{1}{2}\delta_{jk}\text{Re}(\mathbf{D}_k \phi \overline{\mathbf{D}^k \phi})$$

$$+ F_{j\ell} F_k{}^\ell - \frac{1}{4}\delta_{jk} F_{\ell m} F^{\ell m}. \tag{7.4}$$

Given a vector field S on $\mathbb{R}^4$, we define as before the associated 1-and 0-currents

$$^{(S)}J_j[A, \phi] := \mathcal{Q}_{jk}[A, \phi]S^k,$$
$$^{(S)}K[A, \phi] := \mathcal{Q}_{jk}[A, \phi]^{(S)}\pi^{jk}$$

which, thanks to (7.3), satisfy the divergence identity

$$\nabla^{\mathbf{a}}(^{(S)}J[A, \phi]_{\mathbf{a}}) = {}^{(S)}K[A, \phi]. \tag{7.5}$$

Choosing S to be the scaling vector field on $\mathbb{R}^4$ so that, in the rectilinear coordinates

$$S^k = x^k, \quad {}^{(S)}\pi^{jk} = 2\delta^{jk},$$

we have

$${}^{(S)}K[A, \phi] = -2|\mathbf{D}\phi|^2, \quad |{}^{(S)}J_j[A, \phi]| \lesssim |x||\mathbf{D}\phi|^2 + |x||F|^2.$$

where $|\mathbf{D}\phi|^2 = \sum_{j=1}^{4} |\mathbf{D}_j\phi|^2$ and $|F|^2 = \sum_{1 \le j < k \le 4} |F_{jk}|^2$.

We now integrate (7.5) by parts on a ball $B_R \subseteq \mathbb{R}^4$ of radius $R > 1$ centered at 0. Then we see that

$$-2 \int_{B_R} |\mathbf{D}\phi|^2 \, dx = \int_{\partial B_R} {}^{(S)}J[A, \phi]_{\mathbf{a}}\mathbf{n}^{\mathbf{a}}, \quad \text{where } \mathbf{n} = \frac{x^\ell}{|x|}\partial_\ell. \quad (7.6)$$

By the finite energy condition, we have $|\mathbf{D}\phi|, |F| \in L^2(\mathbb{R}^4)$; this fact is enough to deduce the existence of a sequence of radii $R_n \to \infty$ along which the boundary integral vanishes. Hence it follows that $\mathbf{D}_x\phi = 0$.

It only remains to show that $F = 0$. Note that F is now a harmonic 2-form in $L^2(\mathbb{R}^4)$, as $dF = d^2A = 0$ and the right-hand side of the first equation in (7.3) vanishes. Therefore, each component F_{jk} is a harmonic function. By the non-existence[14] of nontrivial harmonic functions in $L^2(\mathbb{R}^4)$, it follows that $F = 0$, which completes the proof. □

7.2 Triviality of finite energy self-similar solutions

In the case of a self-similar solution with finite energy, our main result is as follows.

Proposition 7.2 *Let (A, ϕ) be a smooth solution to ([MKG]) on the forward light cone $C_{(0,\infty)}$ with $\iota_{X_0}F = 0$ and $\mathbf{D}_{X_0}\phi + \frac{1}{\rho}\phi = 0$. Suppose furthermore that (A, ϕ) has finite energy, i.e., $\sup_{t \in (0,\infty)} \mathcal{E}_{S_t}[A, \phi] < \infty$. Then $\mathcal{E}_{S_t}[A, \phi] = 0$ for all $t > 0$.*

Proof We use the hyperbolic coordinates (ρ, y, Θ), in which $X_0 = \partial_\rho$. By the self-similarity assumption $\iota_{X_0}F(\cdot) = F(\partial_\rho, \cdot) = 0$ and $\mathbf{D}_{\partial_\rho}\phi = -\frac{1}{\rho}\phi$, it follows that the pullback of (A, ϕ) to $\mathcal{H}_1 = \{\rho = 1\} = \mathbb{H}^4$, which we still denote by (A, ϕ), solves the system

$$\begin{cases} -\operatorname{div}_{\mathbb{H}^4} F = \operatorname{Im}(\phi\overline{\mathbf{D}_{\mathbb{H}^4}\phi}), \\ (-\triangle_{\mathbb{H}^4, A} - 2)\phi = 0, \end{cases} \quad (7.7)$$

---

[14] This fact can be proved using the monotonicity (7.14), which holds for all $0 < r_1 < r_2$ for harmonic functions on $\mathbb{R}^4$.

where $F = dA$, $(\text{div}_{\mathbb{H}^4} F)_\mathbf{a} = \nabla^\mathbf{b}_{\mathbb{H}^4} F_{\mathbf{ba}}$, $\mathbf{D}_{\mathbb{H}^4} = \nabla_{\mathbb{H}^4} + iA$ and $\triangle_{\mathbb{H}^4, A} = \mathbf{D}^\mathbf{a}_{\mathbb{H}^4} \mathbf{D}_{\mathbb{H}^4, \mathbf{a}}$. Furthermore, by Proposition 5.1 applied to $\mathcal{H}_1 = \mathbb{H}^4$, we have

$$\int_{\mathbb{H}^4} \frac{1}{2} \cosh y \, |F|^2_{\mathbb{H}^4} \, d\sigma_{\mathbb{H}^4} < \infty, \tag{7.8}$$

$$\int_{\mathbb{H}^4} \frac{1}{2} \Big[\cosh y |\phi|^2 + 2 \sinh y \text{Re}[\phi \overline{\mathbf{D}_y \phi}] + \cosh y |\mathbf{D}\phi|^2_{\mathbb{H}^4} \Big] d\sigma_{\mathbb{H}^4} < \infty. \tag{7.9}$$

where $|F|^2_{\mathbb{H}^4} = \frac{1}{2}(g^{-1}_{\mathbb{H}^4})^\mathbf{ac}(g^{-1}_{\mathbb{H}^4})^\mathbf{bd} F_\mathbf{ab} F_\mathbf{cd}$ and $|\mathbf{D}\phi|^2_{\mathbb{H}^4} = (g^{-1}_{\mathbb{H}^4})^\mathbf{ab} \mathbf{D}_\mathbf{a} \phi \overline{\mathbf{D}_\mathbf{b} \phi}$.

In order to proceed, we reformulate the system on $\mathbb{D}^4$ using the conformal equivalence of $\mathbb{D}^4$ and $\mathbb{H}^4$. Consider the following map from $\mathbb{D}^4$ to $\mathbb{H}^4$:

$$\Phi : \mathbb{D}^4 \to \mathbb{H}^4, \quad (r, \Theta) \mapsto (y, \Theta) = \big(2 \tanh^{-1} r, \Theta\big)$$

The map Φ is a conformal isometry, i.e.,

$$\Phi^* g_{\mathbb{H}^4} = \Phi^* \big(dy^2 + \sinh^2 y \, g_{\mathbb{S}^3}\big) = \Omega^2 \big(dr^2 + r^2 \, g_{\mathbb{S}^3}\big) = \Omega^2 g_{\mathbb{D}^4},$$

where Φ^* denotes the pullback along Φ to $\mathbb{D}^4$, and $\Omega := \frac{2}{1-r^2}$. For the pulled-back pair $(\Phi^* A, \Omega \, \Phi^* \phi)$ on $\mathbb{D}^4$, which (slightly abusing the notation) we will denote by (A, u), we have

$$\begin{cases} \partial^\ell F_{j\ell} = \text{Im}(u \overline{\mathbf{D}_j u}) \\ \mathbf{D}^\ell \mathbf{D}_\ell u = 0. \end{cases} \tag{7.10}$$

where $F = dA$ and $\mathbf{D} = \nabla + iA$. Moreover, the bounds (7.8) and (7.9) then translate to

$$\int_{\mathbb{D}^4} \frac{1}{2} \frac{1+r^2}{1-r^2} |F|^2_{\mathbb{D}^4} \, d\sigma_{\mathbb{D}^4} < \infty, \tag{7.11}$$

$$\int_{\mathbb{D}^4} \frac{1}{2} \Big[\frac{1}{1-r^2} |r\mathbf{D}_r u + 2u|^2 + \frac{1}{1-r^2} |\mathbf{D}_r u|^2$$

$$+ \frac{1+r^2}{(1-r^2)r^2} |\slashed{\mathbf{D}} u|^2 \Big] d\sigma_{\mathbb{D}^4} < \infty. \tag{7.12}$$

where $|\slashed{\mathbf{D}} u|^2 = (g^{-1}_{\mathbb{S}^3})^\mathbf{ab} \mathbf{D}_\mathbf{a} u \overline{\mathbf{D}_\mathbf{b} u}$. Indeed, note that

$$\Phi^* d\sigma_{\mathbb{H}^4} = \Omega^4 d\sigma_{\mathbb{D}^4}, \quad \Phi^*(\cosh y) = \frac{1+r^2}{1-r^2}, \quad \Phi^*(\sinh y) = \frac{2r}{1-r^2}.$$

From these identities and (7.8), we immediately see that (7.11) holds. Moreover, (7.12) follows from (7.9) and the following computation:

$$\int_{\mathbb{H}^4} \frac{1}{2} \Big[\cosh y |\phi|^2 + 2 \sinh y \mathrm{Re}[\phi\overline{\mathbf{D}_y\phi}] + \cosh y |\mathbf{D}\phi|^2_{\mathbb{H}^4} \Big] d\sigma_{\mathbb{H}^4}$$

$$= \int_{\mathbb{D}^4} \frac{1}{2} \Big[\frac{1+r^2}{1-r^2} \Omega^2 |u|^2 + \frac{4r}{1-r^2} \mathrm{Re}[\Omega u \overline{\Omega \mathbf{D}_r (\Omega^{-1} u)}]$$

$$+ \frac{1+r^2}{1-r^2} \Big(|\Omega \mathbf{D}_r (\Omega^{-1} u)|^2 + \frac{1}{r^2} |\mathbf{D}u|^2 \Big) \Big] d\sigma_{\mathbb{D}^4}$$

$$= \int_{\mathbb{D}^4} \frac{1}{2} \Big[\frac{4}{1-r^2} |u|^2 + \frac{4r}{1-r^2} \mathrm{Re}[u \overline{\mathbf{D}_r u}] + \frac{r^2+1}{1-r^2} |\mathbf{D}_r u|^2$$

$$+ \frac{1+r^2}{(1-r^2)r^2} |\mathbf{D}u|^2 \Big] d\sigma_{\mathbb{D}^4}$$

$$= \int_{\mathbb{D}^4} \frac{1}{2} \Big[\frac{1}{1-r^2} |r\mathbf{D}_r u + 2u|^2 + \frac{1}{1-r^2} |\mathbf{D}_r u|^2$$

$$+ \frac{1+r^2}{(1-r^2)r^2} |\mathbf{D}u|^2 \Big] d\sigma_{\mathbb{D}^4}.$$

We will now show that (7.10), (7.11) and (7.12) imply $u = 0$ on $\mathbb{D}^4$. Since the system (7.10) coincides with (7.3) restricted to $\mathbb{D}^4$, the divergence identity (7.5) can be used in the present context as well. Integrating (7.5) by parts on a ball $B_R \subseteq \mathbb{D}^4$ of radius $R < 1$ centered at 0, we see that

$$-2\int_{B_R} |\mathbf{D}u|^2 \, d\sigma_{\mathbb{D}^4} = \int_{\partial B_R} {}^{(S)}J[A, u]_{\mathbf{a}} \mathbf{n}^{\mathbf{a}}, \quad \text{where } \mathbf{n} = \frac{x^\ell}{|x|} \partial_\ell. \quad (7.13)$$

Observe that (7.11) and (7.12) imply the existence of a sequence $R_n \to 1$ such that

$$\int_{\partial B_{R_n}} |{}^{(S)}J[A, u]_{\mathbf{a}} \mathbf{n}^{\mathbf{a}}| \to 0,$$

which shows that $\mathbf{D}u = 0$ on $\mathbb{D}^4$. Plugging this information into (7.12), it follows that $u = 0$ on $\mathbb{D}^4$, as desired.

To complete the proof, it only remains to show that $F = 0$. As before, F is now a harmonic 2-form in $L^2(\mathbb{D}^4)$ by (7.7); hence each component F_{jk} is a harmonic function on $\mathbb{D}^4$. Fix $j, k \in \{1, 2, 3, 4\}$ and observe that $\varphi := F_{jk}$, viewed as a real-valued function, obeys the following monotonicity property:

$$\frac{1}{r_1^3} \int_{\partial B_{r_1}} |\varphi|^2 \leq \frac{1}{r_2^3} \int_{\partial B_{r_2}} |\varphi|^2 \quad \text{where } 0 < r_1 < r_2 < 1. \quad (7.14)$$

Indeed, (7.14) is a consequence of interpolating the inequalities

$$\frac{1}{r_1^3} \int_{\partial B_{r_1}} |\varphi| \leq \frac{1}{r_2^3} \int_{\partial B_{r_2}} |\varphi|, \quad \sup_{\partial B_{r_1}} |\varphi| \leq \sup_{\partial B_{r_2}} |\varphi| \quad \text{where } 0 < r_1 < r_2 < 1,$$

which follow from the mean-value property and the weak maximum principle for the subharmonic function $|\varphi|$ on $\mathbb{D}^4$, respectively. By (7.11), it follows that $F_{jk} = \varphi = 0$ on $\mathbb{D}^4$. □

7.3 Regularity of stationary and self-similar weak solutions to (MKG)

We end this section with a regularity result, which applies to weak solutions obtained by Proposition 6.1.

Proposition 7.3 *Let (A, ϕ) be a weak solution to (MKG) on an open set $\mathcal{O} \subseteq \mathbb{R}^{1+4}$ such that*

$$A_\mu \in H^1_{t,x}(\mathcal{O}), \quad \phi \in H^{\frac{3}{2}}_{t,x}(\mathcal{O}). \tag{7.15}$$

Suppose furthermore that one of the following holds:

(1) *Either (A, ϕ) is stationary on $\mathcal{O}$ in the sense of (7.1); or*
(2) *The set $\mathcal{O}$ is a subset of the cone $C_{(0,\infty)} = \{0 \leq r < t\}$ and (A, ϕ) is self-similar on $\mathcal{O}$ in the sense of (7.2).*

Then for every $p \in \mathcal{O}$, there exists an open neighborhood $p \in Q_p \subseteq \mathcal{O}$ and a gauge transformation $\chi_{[p]} \in \mathcal{Y}^w(Q_p)$ such that $(A_{[p]}, \phi_{[p]}) = (A - d\chi_{[p]}, \phi e^{i\chi_{[p]}})$ is smooth on Q_p.

Proof The idea is to derive an elliptic system as in (7.3) [resp. (7.7)] using stationarity [resp. self-similarity], and then use its regularity theory. To get rid of the non-local operator $\langle D_{t,x} \rangle^{\frac{3}{2}}$ in the norm, we begin with the following simple maneuver: For any open bounded subset $Q \subseteq \mathcal{O}$ with smooth boundary, by Sobolev and (7.15), we have

$$A_\mu \in H^1_{t,x}(Q), \quad \phi \in W^{1,q}_{t,x}(Q) \tag{7.16}$$

where $q = \frac{5}{2}$. The important point is that $q > 2$, which will make this bound subcritical. Hence we would be able to conclude regularity via a simple elliptic bootstrap argument.

We first treat Case 1. Applying a suitable Lorentz transformation, it suffices to consider the case $Y = \partial_t$ in the rectilinear coordinates ($t = x^0, x^1, \ldots, x^4$). Moreover, applying an appropriate space-time translation, we may assume that

p is the origin. Let $Q_p := (-\delta, \delta) \times \delta B$, where δB is the open ball of radius δ centered at the origin. Choosing $\delta > 0$ small enough, we have $Q_p \subseteq \mathcal{O}$. By (7.16) and Fubini, there exists $\bar{t} \in (-\delta, \delta)$ such that

$$A \restriction_{\bar{t} \times \delta B} \in H^1(\delta B), \quad \phi \restriction_{\bar{t} \times \delta B} \in W^{1,q}(\delta B), \tag{7.17}$$

where the shorthand $\bar{t} = \{\bar{t}\}$ is used for simplicity. We claim that there exists $\chi_{[p]} \in \mathcal{Y}^w((-\delta, \delta) \times \delta B)$ so that $\chi_{[p]} \restriction_{\bar{t} \times \delta B} \in H^2(\delta B)$ and

$$\partial_t \chi_{[p]} = A_0 \text{ in } (-\delta, \delta) \times \delta B, \quad \triangle \chi_{[p]} \restriction_{\bar{t} \times \delta B} = \partial^\ell (A \restriction_{\bar{t} \times \delta B})_\ell. \tag{7.18}$$

Indeed, we may simply define $\underline{\chi}_{[p]} = \triangle^{-1} \partial^\ell (\eta A \restriction_{\{t = \bar{t}\}})_\ell$, where $\eta \in C_0^\infty(\mathbb{R}^4)$ satisfies $\eta = 1$ on δB and $\operatorname{supp} \eta \subseteq \mathcal{O}$, then solve the transport equation $\partial_t \chi_{[p]} = A_0$ in $(-\delta, \delta) \times \delta B$ with initial data $\chi_{[p]} \restriction_{\bar{t} \times \delta B} = \underline{\chi}_{[p]}$. That this $\chi_{[p]}$ belongs to $\mathcal{Y}^w((-\delta, \delta) \times \delta B)$ and $\chi_{[p]} \restriction_{\bar{t} \times \delta B} \in H^2(\delta B)$ easily follow from the bounds for A in (7.16) and (7.17).

Consider now the gauge transform $(A_{[p]}, \phi_{[p]}) = (A - d\chi_{[p]}, \phi e^{i\chi_{[p]}})$. By (7.18), we have

$$A_{[p]0} = 0 \text{ in } (-\delta, \delta) \times \delta B, \quad \partial^\ell \left(A_{[p]} \restriction_{\bar{t} \times \delta B} \right)_\ell = 0 \text{ in } \delta B. \tag{7.19}$$

By the stationarity assumption $\iota_{\partial_t} F = 0$ and $\mathbf{D}_{\partial_t} \phi = 0$, it follows that

$$\partial_t A_{[p]j} = F_{0j} = 0, \quad \partial_t \phi_{[p]} = 0 \text{ in } (-\delta, \delta) \times \delta B.$$

Hence to prove that $(A_{[p]}, \phi_{[p]})$ is smooth in Q_p, it suffices to show that $(A_{[p]}, \phi_{[p]}) \restriction_{\bar{t} \times \delta B}$ is smooth. Abusing the notation slightly for simplicity, we will henceforth write $A = A_{[p]} \restriction_{\bar{t} \times \delta B}$ and $\phi = \phi_{[p]} \restriction_{\bar{t} \times \delta B}$. By (7.3) and (7.19) (in particular, the Coulomb condition for A), (A, ϕ) satisfies an elliptic system on δB of the schematic form

$$\triangle A = \phi \partial \phi + \phi A \phi,$$
$$\triangle \phi = A \partial \phi + A A \phi.$$

Moreover, (A, ϕ) belongs to $A \in H^1(\delta B)$ and $\phi \in W^{1,q}(\delta B)$, thanks to (7.17) and $\chi_{[p]} \restriction_{\bar{t} \times \delta B} \in H^2(\delta B)$. As this system is H^1-critical and every nonlinear term has at least one factor of ϕ, which obeys a subcritical bound $\phi \in W^{1,q}(\delta B)$, we can perform a standard elliptic bootstrap argument to conclude that (A, ϕ) is smooth on δB with uniform bounds on compact subsets. This concludes the proof in Case 1.

The proof in Case 2 is entirely analogous to Case 1, so we only give a brief outline. Here, instead of the rectilinear coordinates, we use the hyperbolic

coordinates (ρ, y, Θ), in which $X = \partial_\rho$. Applying a suitable Lorentz transformation and scaling transformation, we may assume that p coincides with the point $\rho = 1$, $y = 0$. Let $Q_p = (-\delta, \delta) \times D_\delta$, where $D_\delta := \{(y, \Theta) : |y| < \delta\}$, which is contained in $\mathcal{O}$ if $\delta > 0$ is sufficiently small. By (7.16) and Fubini, there exists $\overline{\rho} \in (-\delta, \delta)$ such that

$$A \upharpoonright_{\overline{\rho} \times D_\delta} \in H^1(D_\delta), \quad \phi \upharpoonright_{\overline{\rho} \times D_\delta} \in W^{1,p}(D_\delta). \tag{7.20}$$

Proceeding as before, we can find $\chi_{[p]} \in \mathcal{Y}^w((-\delta, \delta) \times D_\delta)$ so that $\chi_{[p]} \upharpoonright_{\overline{\rho} \times D_\delta} \in H^2(D_\delta)$ and

$$\partial_\rho \chi_{[p]} = 0 \quad \text{in } (-\delta, \delta) \times D_\delta, \quad \triangle_{\mathcal{H}_{\overline{\rho}}} \chi_{[p]} \upharpoonright_{\overline{\rho} \times D_\delta} = \nabla^{\mathbf{a}}_{\mathcal{H}_{\overline{\rho}}} \left(A \upharpoonright_{\overline{\rho} \times D_\delta} \right)_{\mathbf{a}}.$$

Then the gauge transform $(A_{[p]}, \phi_{[p]}) = (A - d\chi_{[p]}, \phi e^{i\chi_{[p]}})$ obeys

$$A_{[p]}(\partial_\rho) = 0 \text{ in } (-\delta, \delta) \times D_\delta, \quad \nabla^{\mathbf{a}}_{\mathcal{H}_{\overline{\rho}}} \left(A_{[p]} \upharpoonright_{\overline{\rho} \times D_\delta} \right)_{\mathbf{a}} = 0 \text{ in } D_\delta.$$

By self-similarity, we have $\mathcal{L}_{\partial_\rho} A_{[p]} = 0$ and $\partial_\rho(\rho\phi_{[p]}) = 0$, so it only remains to prove that the pullback of $(A_{[p]}, \phi_{[p]})$ on $\overline{\rho} \times D_\delta$, which we will refer to as (A, ϕ), is smooth. As in the previous case, this is a consequence of the fact that (A, ϕ) obeys an elliptic system (thanks to (7.7) and the Coulomb gauge condition on $\mathcal{H}_{\overline{\rho}}$), the bounds $A \in H^1(D_\delta)$ and $\phi \in W^{1,q}(D_\delta)$ with $q > 2$ (by (7.20) and $\chi_{[p]} \upharpoonright_{\overline{\rho} \times D_\delta} \in H^2(D_\delta)$), and a standard elliptic bootstrap argument.

□

8 Proof of global well-posedness and scattering

Here we carry out the proof of Theorem 1.3 using the tools developed in the earlier parts.

8.1 Finite time blow-up/non-scattering scenarios and initial reduction

Our overall strategy for proving Theorem 1.3 is by contradiction. Suppose that Theorem 1.3 fails for an initial data set $(a, e, f, g) \in \mathcal{H}^1$ in the global Coulomb gauge. By time reversal symmetry, it suffices to consider the forward evolution. Let (A, ϕ) be the admissible $C_t \mathcal{H}^1$ solution to the Cauchy problem in the global Coulomb gauge defined on the maximal forward time interval $I = [0, T_+)$ for some $T_+ > 0$ constructed by Theorem 4.3. By Theorem 4.8, the solution (A, ϕ) exhibits one of the following behaviors:

(1) *Finite time blow-up* We have $T_+ < \infty$ and

$$\|A_0\|_{Y^1[0,T_+)} + \|A_x\|_{S^1[0,T_+)} + \|\phi\|_{S^1[0,T_+)} = \infty. \tag{8.1}$$

(2) *Non-scattering* We have $T_+ = \infty$, but

$$\|A_0\|_{Y^1[0,\infty)} + \|A_x\|_{S^1[0,\infty)} + \|\phi\|_{S^1[0,\infty)} = \infty. \tag{8.2}$$

In the case of finite time blow-up, we may use the energy concentration scale r_c in Theorem 4.3 to show that the energy must concentrate at a point.

Lemma 8.1 *Let (A, ϕ) be an admissible $C_t\mathcal{H}^1$ solution to (MKG) on $[0, T_+) \times \mathbb{R}^4$ with $T_+ < \infty$ in the global Coulomb gauge. Then either (A, ϕ) can be continued past T_+ as an admissible $C_t\mathcal{H}^1$ solution in the global Coulomb gauge (as in Theorem 4.3), or there exists a point $x_0 \in \mathbb{R}^4$ such that*

$$\limsup_{t \to T_+} \mathcal{E}_{\{t\} \times B_{(T_+ - t)}(x_0)}[A, \phi] > 0. \tag{8.3}$$

Proof For $t < T_+$ and $x \in \mathbb{R}^4$ we define the function

$$E(t, x) = \mathcal{E}_{\{t_0\} \times B_{(T_+ - t)}(x)}[A, \phi].$$

This is continuous in x, and, by the non-negativity of the flux in the energy relation (5.3), it is non-increasing in t. Further, by the same relation, we have

$$\lim_{|x| \to \infty} E(t, x) = 0, \qquad \text{uniformly in } t \in [0, T_+). \tag{8.4}$$

Then we have two alternatives:

(i) Either $\lim_{t \to T_+} \sup_{x \in \mathbb{R}^4} E(t, x) < \delta_0(E, \epsilon_*^2)$, which implies that there exists t_0 so that energy concentration scale r_c at $t = t_0$ as in (4.4) is greater than $T_+ - t_0$. By Theorem 4.3 we can then extend (A, ϕ) past T_+, as claimed.

(ii) Or, $\lim_{t \to T_+} \sup_{x \in \mathbb{R}^4} E(t, x) \geq \delta_0(E, \epsilon_*^2)$. Then the sets $D_t = \{x \in \mathbb{R}^4 : E(t, x) \geq \frac{1}{2}\delta_0(E, \epsilon_*^2)\}$ are nonempty and decreasing in t. Moreover, they are compact by (8.4). Thus they must intersect. Any x_0 in the intersection will provide the second alternative in the lemma. $\square$

Theorem 4.7 provides additional information about the nature of the singularity in both scenarios, which is crucial to our proof of Theorem 1.3. To utilize this information, we introduce a smooth function ζ satisfying the following properties:

- $\operatorname{supp} \zeta \subseteq B_1(0)$ and $\int \zeta = 1$.
- There exists a function $\widetilde{\zeta} \in C_0^\infty(\mathbb{R}^4)$ with $\widetilde{\zeta} \geq 0$ such that $\zeta = \widetilde{\zeta} * \widetilde{\zeta}$.

Then we define the physical space version of energy dispersion as follows:

$$\mathbf{ED}[A, \phi](I) := \sup_{k \in \mathbb{Z}} \left(2^{-k} \|\zeta_{2^{-k}} * \phi(t, x)\|_{L_{t,x}^\infty(I \times \mathbb{R}^4)} + 2^{-2k} \|\zeta_{2^{-k}} * \mathbf{D}_t \phi(t, x)\|_{L_{t,x}^\infty(I \times \mathbb{R}^4)} \right)$$

$$(8.5)$$

where $\zeta_{2^{-k}} := 2^{4k} \zeta(2^k \cdot)$, and the convolution $*$ is defined with respect to only the spatial variables $(x^1, \ldots, x^4)$. The first property makes $\mathbf{ED}[A, \phi]$ simpler to use in physical space arguments; on the other hand, the second property is helpful in connection with the diamagnetic inequality, which we state here.

Lemma 8.2 (Diamagnetic inequality) *Let $O \subseteq \mathbb{R}^4$ be an open set and $\phi, A \in H^1(O)$. Then for any smooth vector X, $|\partial_X |\phi|| \leq |\mathbf{D}_X \phi|$ in the sense of distributions. More precisely, for any smooth $\eta \geq 0$ with $\operatorname{supp} \eta \subseteq O$, we have*

$$\int \eta |\partial_X |\phi|| \, \mathrm{d}x \leq \int \eta |\mathbf{D}_X \phi| \, \mathrm{d}x.$$

The key to the proof is the formal computation $|\partial_X |\phi|| = ||\phi|^{-1} \langle \phi, \mathbf{D}_X \phi \rangle| \leq |\mathbf{D}_X \phi|$; we omit the standard details. We fix the choice of functions $\zeta, \tilde{\zeta}$ here, and henceforth we will suppress the dependence of constants on these functions for simplicity.

The physical space version $\mathbf{ED}[A, \phi]$ is related to the earlier Littlewood-Paley version $ED[\phi]$ defined in (4.9) as follows.

Lemma 8.3 *Let (A, ϕ) be an admissible $C_t \mathcal{H}^1$ solution to (MKG) on $I \times \mathbb{R}^4$ in the global Coulomb gauge with $\mathcal{E}_{\{t\} \times \mathbb{R}^4}[A, \phi] \leq E$. Then there exists $C = C(E)$ such that*

$$ED[\phi](I) \leq C \, \mathbf{ED}[A, \phi](I) + \frac{1}{100} \epsilon(E),$$

where $\epsilon(E)$ is as in Theorem 4.7.

Proof All norms in this proof will be taken over $I \times \mathbb{R}^4$. The following estimates are straightforward to establish:

$$\sup_k 2^{-k} \|P_k \phi\|_{L_{t,x}^\infty} \lesssim \sup_k 2^{-k} \|\zeta_{2^{-k}} * \phi\|_{L_{t,x}^\infty}, \quad (8.6)$$

$$\sup_k 2^{-2k} \|P_k (\mathbf{D}_t \phi)\|_{L_{t,x}^\infty} \lesssim \sup_k 2^{-2k} \|\zeta_{2^{-k}} * (\mathbf{D}_t \phi)\|_{L_{t,x}^\infty}. \quad (8.7)$$

As these two estimates are proved in the same manner, we only consider (8.6). Fix $k \in \mathbb{Z}$, and let $m_0 > 0$ be an absolute constant to be chosen. We have

$$2^{-k}\|P_k\phi\|_{L_x^\infty} \le 2^{-k}\|\zeta_{2^{-k-m_0}} * P_k\phi\|_{L_x^\infty} + 2^{-k}\|(1-\zeta_{2^{-k-m_0}}*)P_k\phi\|_{L_x^\infty}$$

$$\lesssim 2^{m_0}2^{-k-m_0}\|\zeta_{2^{-k-m_0}}*\phi\|_{L_x^\infty} + 2^{-m_0}2^{-2k}\sup_{j=1,\dots,4}\|\partial_j P_k\phi\|_{L_x^\infty}$$

The last term can be bounded by $\lesssim 2^{-m_0}2^{-k}\|P_k\phi\|_{L_x^\infty}$, which can be absorbed into the left-hand side by taking m_0 sufficiently large. Hence (8.6) follows.

In view of (8.6) and (8.7), the lemma would follow once we prove that, for any $m_1 > 10$,

$$\sup_k 2^{-2k}\|P_k\partial_t\phi\|_{L_{t,x}^\infty} \lesssim_E 2^{m_1}\sup_k\left(2^{-2k}\|P_k(\mathbf{D}_t\phi)\|_{L_{t,x}^\infty}\right.$$
$$\left. + 2^{-k}\|P_k\phi\|_{L_{t,x}^\infty}\right) + 2^{-m_1}.$$

By the relation $\partial_t = \mathbf{D}_t - iA_0$, it suffices to show that

$$\sup_k 2^{-2k}\|P_k(A_0\phi)\|_{L_{t,x}^\infty} \lesssim 2^{m_1}E^{\frac{1}{2}}\sup_k 2^{-k}\|P_k\phi\|_{L_{t,x}^\infty} + 2^{-m_1}\left(E + E^{\frac{3}{2}}\right).$$
(8.8)

Thanks to the global Coulomb condition, we have

$$\|A_0\|_{L_t^\infty\dot{H}_x^1} \lesssim E^{1/2}, \quad \|\phi\|_{L_t^\infty\dot{H}_x^1} \lesssim E^{1/2} + E.$$

For each $k \in \mathbb{Z}$, we split $\phi = P_{\le k+m_1}\phi + P_{>k+m_1}\phi$. For the former, we have

$$2^{-2k}\|P_k(A_0 P_{\le k+m_1}\phi)\|_{L_{t,x}^\infty} \lesssim \sum_{\ell \le k+m_1} 2^{\ell-k}\|A_0\|_{L_t^\infty L_x^4}2^{-\ell}\|P_\ell\phi\|_{L_{t,x}^\infty}$$

$$\lesssim 2^{m_1}E^{\frac{1}{2}}\sup_\ell 2^{-\ell}\|P_\ell\phi\|_{L_{t,x}^\infty}.$$

For the latter, by the properties of frequency supports, note that

$$P_k(A_0 P_{>k+m_1}\phi) = \sum_{\ell > k+m_1} P_k(P_{[\ell-3,\ell+3]}A_0 P_\ell\phi).$$

Hence (8.8) follows from the estimate

$$2^{-2k}\|P_k(A_0 P_{>k+m_1}\phi)\|_{L_{t,x}^\infty} \lesssim \sum_{\ell > k+m_1} 2^{2k}\|P_{[\ell-3,\ell+3]}A_0\|_{L_t^\infty L_x^2}\|P_\ell\phi\|_{L_t^\infty L_x^2}$$

$$\lesssim 2^{-2m_1}(E + E^{3/2}).$$

$\square$

As a result, there exists a function $\mathbf{e} = \mathbf{e}(E) > 0$ such that Theorem 4.7 holds with the condition (4.11) replaced by

$$\mathbf{ED}[A, \phi](I) \leq \mathbf{e}(E). \tag{4.11'}$$

Let $\varepsilon > 0$ be a small parameter to be chosen below. We have the following result, which unifies the proof of Theorem 1.3 in both finite time blow-up and non-scattering scenarios from here on.

Lemma 8.4 *Suppose that Theorem 1.3 fails for some initial data (a, e, f, g) of energy E. Then for every $\varepsilon > 0$ there exists a sequence $\varepsilon_n \to 0$ and a sequence of admissible $C_t \mathcal{H}^1$ solutions $(A^{(n)}, \phi^{(n)})$ on $[\varepsilon_n, 1] \times \mathbb{R}^4$ in the global Coulomb gauge that satisfy the following properties:*

(1) *Bounded energy in the cone*

$$\mathcal{E}_{S_t}\left[A^{(n)}, \phi^{(n)}\right] \leq 2E \quad \text{for every } t \in [\varepsilon_n, 1], \tag{8.9}$$

(2) *Small energy outside the cone*

$$\mathcal{E}_{(\{t\} \times \mathbb{R}^4) \setminus S_t}\left[A^{(n)}, \phi^{(n)}\right] \leq \varepsilon^8 E \quad \text{for every } t \in [\varepsilon_n, 1], \tag{8.10}$$

(3) *Decaying flux on ∂C*

$$\mathcal{F}_{[\varepsilon_n, 1]}\left[A^{(n)}, \phi^{(n)}\right] + \mathcal{G}_{S_1}\left[\phi^{(n)}\right] \leq \varepsilon_n^{\frac{1}{2}} E, \tag{8.11}$$

(4) *Pointwise concentration at $t = 1$*

$$2^{-k_n} |\zeta_{2^{-k_n}} * \phi^{(n)}(1, x_n)| + 2^{-2k_n} |\zeta_{2^{-k_n}} * \mathbf{D}_t^{(n)} \phi^{(n)}(1, x_n)| > \mathbf{e}(E) \tag{8.12}$$

for some $k_n \in \mathbb{Z}$ and $x_n \in \mathbb{R}^4$.

Remark 8.5 The small parameter $\varepsilon > 0$ will be specified near the end of the proof of Theorem 1.3, precisely in Lemma 8.11, depending only on E.

Remark 8.6 By the global Coulomb gauge condition $\partial^\ell A_\ell^{(n)} = 0$, the following gauge dependent uniform bounds for $A^{(n)}$ and $\phi^{(n)}$ hold:

$$\|\partial_{t,x} A^{(n)}\|_{L_t^\infty([\varepsilon_n, 1]; L_x^2)} \lesssim E^{\frac{1}{2}}, \quad \|\partial_{t,x} \phi^{(n)}\|_{L_t^\infty([\varepsilon_n, 1]; L_x^2)} \lesssim \left(1 + E^{\frac{1}{2}}\right) E^{\frac{1}{2}}. \tag{8.13}$$

Proof Suppose that Theorem 1.3 fails. Then by the discussion at the beginning of the section, there exists an admissible $C_t \mathcal{H}^1$ solution (A, ϕ) of energy E to (MKG) on $[0, T_+) \times \mathbb{R}^4$ which satisfies either $0 < T_+ < \infty$ and (8.1) (finite time blow-up) or $T_+ = \infty$ and (8.2) (non-scattering). We treat these two cases separately.

Case 1: Finite time blow-up By Lemma 8.1, there exists a point $x_0 \in \mathbb{R}^4$ such that (8.3) holds. By translation in space-time and reversing time, we may assume that $x_0 = 0$ and we have energy concentration at the space-time origin as $t \to 0$, i.e.,

$$\limsup_{t \to 0} \mathcal{E}_{S_t}[A, \phi] > 0. \tag{8.14}$$

Our next course of action is to use the excision and gluing technique (Proposition 4.4) to cut away the part of (A, ϕ) outside the cone of influence of $(0, 0)$. In what follows, we denote the ball $B_1(0)$ by B, so that $rB = B_r(0)$ for any $r > 0$.

By Corollary 5.3 there exists $t_0 > 0$ such that

$$\mathcal{F}_{\partial C_{(0, t_0]}}[A, \phi] \ll \min\left\{\delta_0(E, \epsilon_*^2), \varepsilon^8 E\right\}$$

where $\delta_0(E, \epsilon_*^2)$ is as in (4.4). Furthermore, we can find a collar of radius $r_0 > 0$ around $S_{t_0} = \{t_0\} \times t_0 B$ with small energy, i.e.,

$$\mathcal{E}_{\{t_0\} \times ((t_0+r_0)B \setminus t_0 B)}[A, \phi] \ll \min\left\{\delta_0(E, \epsilon_*^2), \varepsilon^8 E\right\}.$$

By local conservation of energy, we then have

$$\mathcal{E}_{\{t\} \times ((t+r_0)B \setminus tB)}[A, \phi] \ll \min\left\{\delta_0\left(E, \epsilon_*^2\right), \varepsilon^8 E\right\} \quad \text{for every } t \in (0, t_0].$$

Observe that the ratio $(t + r_0)/t$ goes to ∞ as $t \to 0$. Hence, by the improved Hardy estimate in Lemma 4.5, for sufficiently small $0 < \bar{t} < r_0$ we also obtain

$$\left\|\frac{1}{|x|}\phi(\bar{t}, \cdot)\right\|^2_{L_x^2(2\bar{t}B \setminus \bar{t}B)} \ll \min\left\{\delta_0\left(E, \epsilon_*^2\right), \varepsilon^8 E\right\} \quad \text{for every } t \in (0, t_0].$$

We may now apply Proposition 4.4 to $(a, e, f, g) = (A_j, F_{0j}, \phi, \mathbf{D}_t\phi)\restriction_{\{t=\bar{t}\}}$ to obtain a new data set $(\widetilde{a}, \widetilde{e}, \widetilde{f}, \widetilde{g})$ that coincides with (a, e, f, g) on $\bar{t}B$ and obeys

$$\mathcal{E}_{\mathbb{R}^4 \setminus \bar{t}B}[\widetilde{a}, \widetilde{e}, \widetilde{f}, \widetilde{g}] \leq \frac{1}{2} \min\left\{\delta_0\left(E, \epsilon_*^2\right), \varepsilon^8 E\right\}.$$

To pass to the global Coulomb gauge, we define the gauge transformation $\chi \in \mathcal{G}^2(\mathbb{R}^4)$ by $\chi = \triangle^{-1} \partial^\ell \widetilde{a}_\ell$ and let $(\check{a}, \check{e}, \check{f}, \check{g})$ be the gauge transform of $(\widetilde{a}, \widetilde{e}, \widetilde{f}, \widetilde{g})$ by χ. Let $(\check{A}, \check{\phi})$ be the admissible $C_t \mathcal{H}^1$ solution to the Cauchy problem in the global Coulomb gauge given by Theorem 4.3, defined on the maximal time interval $I \ni \bar{t}$.

As a consequence of the construction and local conservation of energy, the energy outside the cone C is always tiny, i.e.,

$$\mathcal{E}_{(\{t\} \times \mathbb{R}^4) \setminus S_t} \left[\check{A}, \check{\phi} \right] \leq \frac{1}{2} \min \left\{ \delta_0 \left(E, \epsilon_*^2 \right), \varepsilon^8 E \right\} \quad \text{for every } t \in I. \quad (8.15)$$

Then by an argument similar to the proof of Lemma 8.1, it follows that $(\check{A}, \check{\phi})$ can be always continued to the past until 0, i.e., $(0, \bar{t}] \subseteq I$. Furthermore, there exist sequences $(t_n, x_n) \in I \times \mathbb{R}^4$ and $k_n \in \mathbb{Z}$ with $t_n \to 0$ such that

$$2^{-k_n} |\zeta_{2^{-k_n}} * \check{\phi}(t_n, x_n)| + 2^{-2k_n} |\zeta_{2^{-k_n}} * \check{\mathbf{D}}_t \check{\phi}(t_n, x_n)| > \mathbf{e}(E). \quad (8.16)$$

For otherwise, there exists $\delta > 0$ such that (4.11′) holds on $(0, \delta)$. Then by Theorem 4.7 (with (4.11) replaced by (4.11′)) and Theorem 4.8, the solution $(\check{A}, \check{\phi})$ can be extended past $t = 0$. Hence $\limsup_{t \to 0} \mathcal{E}_{S_t}[\check{A}, \check{\phi}] = 0$, but this fact contradicts (8.14) as $\mathcal{E}_{S_t}[\check{A}, \check{\phi}] = \mathcal{E}_{S_t}[A, \phi]$ for every $t \in I$.

Applying Corollary 5.3 to $(\check{A}, \check{\phi})$, we may choose a sequence $\varepsilon_n \to 0$ such that

$$\mathcal{F}_{[\varepsilon_n t_n, t_n]}[A, \phi] + \mathcal{G}_{S_{t_n}}[\phi] \leq \varepsilon_n^{\frac{1}{2}} E.$$

By the scaling properties of $\mathcal{E}, \mathcal{F}$ and $\mathcal{G}$, as well as scale invariance of (8.16) (which is immediate from definition), it follows that the sequence of rescaled solutions

$$\left(A^{(n)}, \phi^{(n)} \right) (t, x) := t_n \left(\check{A}, \check{\phi} \right) (t_n t, t_n x)$$

obeys the desired properties.

Case 2: Non-scattering This case follows by a simple rescaling argument. Let $R_0 > 0$ be a large radius such that $\mathcal{E}_{\{0\} \times (\mathbb{R}^4 \setminus B_{R_0}(0))}[A, \phi] \leq \varepsilon^8 E$. Translating in time by R_0 and using the local conservation of energy, we may assume that (A, ϕ) obeys

$$\mathcal{E}_{(\{t\} \times \mathbb{R}^4) \setminus S_t}[A, \phi] \leq \varepsilon^8 E \quad \text{for every } t \in [R_0, \infty).$$

By Theorem 4.7 with (4.11) replaced by (4.11′) and (8.2), there exist sequences $(t_n, x_n) \in [R_0, \infty) \times \mathbb{R}^4$ and $k_n \in \mathbb{Z}$ with $t_n \to \infty$ such that

$$2^{-k_n}|\zeta_{2^{-k_n}} * \phi(t_n, x_n)| + 2^{-2k_n}|\zeta_{2^{-k_n}} * \mathbf{D}_t\phi(t_n, x_n)| > \mathbf{e}(E)$$

By Corollary 5.3, we may then choose a sequence $\varepsilon_n \to 0$ such that $\varepsilon_n t_n \to \infty$ and

$$\mathcal{F}_{[\varepsilon_n t_n, t_n]}[A, \phi] + \mathcal{G}_{S_{t_n}}[\phi] \leq \varepsilon_n^{\frac{1}{2}} E.$$

Defining $(A^{(n)}, \phi^{(n)})(t, x) := t_n(A, \phi)(t_n t, t_n x)$, we obtain a desired sequence.

8.2 Elimination of the null concentration scenario

Using Proposition 5.4, in particular the weighted energy estimate on S_1, we show that null concentration cannot happen. The precise statement is as follows.

Lemma 8.7 (No null concentration) *Let* $(A^{(n)}, \phi^{(n)})$ *be a sequence of admissible* $C_t\mathcal{H}^1$ *solutions to* (MKG) *satisfying the conclusions of Lemma* 8.4 *with the sequences* ε_n, k_n *and* x_n. *There exist* $K = K(E) > 0$ *and* $\gamma = \gamma(E) \in (0, 1)$ *such that if* $k_n > K(E)$ *and* $|x_n| > \gamma(E)$ *for all sufficiently large n, and* $\varepsilon > 0$ *is sufficiently small depending on E, then*

$$\limsup_{n \to \infty} 2^{-k_n}|\zeta_{2^{-k_n}} * \phi(1, x_n)| + 2^{-2k_n}|\zeta_{2^{-k_n}} * \mathbf{D}_t^{(n)}\phi^{(n)}(1, x_n)| \leq \mathbf{e}(E).$$
(8.17)

Remark 8.8 Note that $K(E)$ in Lemma 8.7 can be replaced *a posteriori* by any number greater than $K(E)$. Hence given any $m = m(E)$ depending only on E, we may assume in addition to the statement of Lemma 8.7 that

$$2^{-K} \leq \frac{1}{100\, m(E)}(1 - \gamma).$$
(8.18)

This observation will be useful in the proof of Lemma 8.9 below.

Proof The idea of the proof is similar to that of [33, Lemma 6.2] with additional ideas to deal with the presence of covariant derivatives.
Step 1 The starting point is Proposition 5.4 applied to $(A, \phi) = (A^{(n)}, \phi^{(n)})$ with $\varepsilon = \varepsilon_n$, more precisely the first term on the left-hand side of (5.11). Using Lemma 5.10 to write out $^{(X_{\varepsilon_n})}P_T$, we see that the following a-priori estimate holds on S_1:

$$\int_{S_1} \frac{1}{(1 - |x| + \varepsilon_n)^{\frac{1}{2}}} \left(|\mathbf{D}_L^{(n)}\phi^{(n)}|^2 + |\mathbf{D}^{(n)}\phi^{(n)}|^2\right) \mathrm{d}x \lesssim E.$$
(8.19)

By the smallness of the energy outside S_1, we then obtain the global bound

$$\int_{\{t=1\}} \frac{1}{((1-|x|)_+ + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} \left(|\mathbf{D}_L^{(n)} \phi^{(n)}|^2 + |\mathbf{\slashed{D}}^{(n)} \phi^{(n)}|^2 \right) dx \lesssim E,$$

(8.20)

where $(\cdot)_+ := \max\{\cdot, 0\}$.

Step 2 We claim that for any non-negative $k \in \mathbb{Z}$ the following estimate holds:

$$\limsup_{n \to \infty} 2^{-k} |\zeta_{2^{-k}} * |\phi^{(n)}|(1, x)| \lesssim \left(2^{-\frac{3}{8}k} + \left((1-|x|)_+ + 2^{-k}\right)^{\frac{1}{4}} + \varepsilon^4 \right) E^{\frac{1}{2}}.$$

(8.21)

The point of (8.21) is that $|\phi^{(n)}|$ is gauge invariant, and hence we can avoid estimating A. Henceforth, we will denote $\psi^{(n)} := |\phi^{(n)}|(1, \cdot)$. We use the rotational symmetry to bring x to the x^1-axis, so that $x = (|x|, 0, 0, 0)$. Henceforth we will write $x = (x^1, x')$ where $x' = (x^2, x^3, x^4)$.

By the diamagnetic inequality (Lemma 8.2), conservation of energy implies

$$\int |\nabla \psi^{(n)}|^2 \, dx \lesssim E.$$

(8.22)

where $|\nabla \psi|^2 := \sum_{\ell=1}^4 |\partial_\ell \psi|^2$. Note that (8.22) and Young's inequality implies the trivial bound $2^{-k} \|\zeta_{2^{-k}} * \psi^{(n)}\|_{L_x^\infty} \lesssim E^{1/2}$, which allows us to restrict our attention to $x = (x^1, 0, 0, 0)$ with $1/2 < x^1 < 2$.

We claim that for n sufficiently large so that $\varepsilon_n^{1/2} \leq \frac{1}{100} \varepsilon^8$, the directional derivatives other than ∂_1 obey an improved estimate

$$\sum_{j=1}^4 \int w_k |\partial_j \psi^{(n)}|^2 \, dx \lesssim E,$$

(8.23)

where $w_k > 0$ is defined as

$$w_k(x) := \frac{1}{(|1-x^1| + |x'|^2 + 2^{-k})^{\frac{1}{2}} + \varepsilon^8}.$$

(8.24)

To prove (8.23) under the assumption $\varepsilon_n^{1/2} \leq \frac{1}{100} \varepsilon^8$, it suffices to prove

$$\sum_{j=2}^4 \int \frac{1}{(|1-x^1| + |x'|^2 + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} |\partial_j \psi^{(n)}|^2 \, dx \lesssim E.$$

(8.25)

The estimate (8.25) is a consequence of (8.20). Indeed, the latter estimate combined with the diamagnetic inequality implies

$$\int \frac{1}{((1 - |x|)_+ + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} |\slashed{\nabla} \psi^{(n)}|^2 \, dx \lesssim E. \tag{8.26}$$

At $x = (1, 0, 0, 0)$ we have $\frac{1}{r^2} g_{\mathbb{S}^3}^{-1} = \sum_{j=2}^{4} \partial_j \cdot \partial_j$. Therefore, by smoothness, we have

$$\left| |\slashed{\nabla} \psi|^2 - \sum_{j=2}^{4} |\partial_j \psi|^2 \right| \lesssim \left(|1 - x^1| + |x'| \right) |\nabla \psi|^2.$$

On the other hand, $(1 - |x|)_+ \lesssim |1 - x^1| + |x'|^2$.

Combining these facts, we may now bound the left-hand side of (8.25) by

$$\int \frac{1}{(|1 - x^1| + |x'|^2 + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} \left| \sum_{j=2}^{4} |\partial_j \psi^{(n)}|^2 - |\slashed{\nabla} \psi|^2 \right|$$

$$+ \frac{1}{((1 - |x|)_+ + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} |\slashed{\nabla} \psi^{(n)}|^2 \, dx$$

$$\lesssim \int \frac{|1 - x^1| + |x'|}{(|1 - x^1| + |x'|^2 + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} |\nabla \psi|^2 \, dx$$

$$+ \int \frac{1}{((1 - |x|)_+ + \varepsilon_n)^{\frac{1}{2}} + \varepsilon^8} |\slashed{\nabla} \psi^{(n)}|^2 \, dx.$$

Then using (8.22) and (8.26), the desired estimate (8.25) follows.

Compared to the weight in the preceding expression, observe that we have absorbed ε_n into ε^8 and added 2^{-k} in w_k. This maneuver ensures that w_k is slowly varying at scale $2^{-k} \times 2^{-k/2} \times \cdots \times 2^{-k/2}$, i.e., for any $x, y \in \mathbb{R}^4$ we have

$$\left| \frac{w_k(x)}{w_k(x - y)} \right| \lesssim e^{\sum_{j=1}^{4} |y^j| \|\partial_j \log w_k\|_{L^\infty}} \lesssim e^{2^k |y^1| + 2^{k/2} |y'|}. \tag{8.27}$$

We now turn to the task of deriving (8.21) from (8.22) and (8.23). We introduce the notation $Z_k \psi := \zeta_{2^{-k}} * \psi$ and write $z_k(\xi)$ for the symbol of the integral operator Z_k; of course, z_k is nothing but the (spatial) Fourier transform of $\zeta_{2^{-k}}$. Given a smooth cut-off η on $\mathbb{R}^3$ adapted to the unit ball, we furthermore decompose

$$Z_k = Z_k^1 \partial_1 + Z_k^2 \partial_2 + \cdots + Z_k^4 \partial_4$$

where the symbols $z_k^j(\xi)$ of Z_k^j are given by

$$z_k^1(\xi) = z_k(\xi)\eta\left(2^{-\frac{k}{2}}\xi'\right)\frac{1}{i\xi_1},$$

$$z_k^j(\xi) = z_k(\xi)\left(1 - \eta\left(2^{-\frac{k}{2}}\xi'\right)\right)\frac{\xi_j}{i|\xi'|^2} \quad \text{for } j = 2, 3, 4,$$

where $\xi' = (\xi_2, \xi_3, \xi_4)$.

The contribution of $Z_k^1\partial_1$ to (8.21) is easy to treat. Observe that $z_k^1(\xi)i\xi_1$ is a smooth symbol which is rapidly decaying at scale 2^k in the ξ_1-direction, i.e., for every $N \geq 0$ we have

$$|(\xi_1\partial_{\xi_1})^N\left(z_k^1 i\xi_1\right)| \lesssim_N \left(1 + 2^{-k}|\xi_1|\right)^{-100}.$$

Moreover, $z_k^1(\xi)i\xi_1$ is compactly supported in the set $\{|\xi'| \lesssim 2^{k/2}\}$ in the other directions. For any $N \geq 0$, these facts immediately imply the kernel bound

$$|\mathcal{F}_x^{-1}(z_k^1 i\xi_1)| \lesssim_N 2^{\frac{5}{2}k}(1 + 2^k|x^1|)^{-N}(1 + 2^{\frac{k}{2}}|x'|)^{-N},$$

where $\mathcal{F}_x$ denotes the inverse (spatial) Fourier transform. In particular, $\|\mathcal{F}_x^{-1}(z_k^1 i\xi_1)\|_{L_x^{\frac{4}{3}}} \lesssim 2^{\frac{5}{8}k}$. By Young's inequality, we have

$$2^{-k}|Z_k^1\partial_1\psi^{(n)}(x)| \lesssim 2^{-\frac{3}{8}k}\|\psi^{(n)}\|_{\dot{H}_x^1} \lesssim 2^{-\frac{3}{8}k}E^{\frac{1}{2}},$$

which is acceptable.

It remains to treat the contribution of $Z_k^j\partial_j$ for $j = 2, 3, 4$. Denote by $\zeta_k^j(x)$ the integral kernel of Z_k^j, which is simply the inverse Fourier transform of z_k^j. A straightforward computation shows that $\|z_k^j\|_{L_\xi^2} \lesssim 2^k$. Therefore, by Plancherel,

$$\|\zeta_k^j\|_{L_x^2} \lesssim 2^k. \tag{8.28}$$

Next, for any $N, M_2, M_3, M_4 \geq 0$, we have

$$\left|(\xi_1\partial_{\xi_1})^N(\xi_2\partial_{\xi_2})^{M_2}(\xi_3\partial_{\xi_3})^{M_3}(\xi_4\partial_{\xi_4})^{M_4}\left(\sum_{j=2}^4 i\xi_j z_k^j\right)\right|$$

$$\lesssim_{N,M_2,M_3,M_4}\left(1 + 2^{-k}|\xi|\right)^{-100}.$$

In addition, the left-hand side is supported in $\{|\xi'| \lesssim 2^{k/2}\}$ if $M_2 + M_3 + M_4 \neq 0$. Taking the inverse Fourier transform, we obtain

$$\left| \sum_{j=2}^{4} \partial_j \zeta_k^j(x) \right| \lesssim_N \min \left\{ 2^{\frac{3}{2}k}, \left(2^{\frac{k}{2}} |x'| \right)^{-N} \right\} \left(1 + 2^k |x^1| \right)^{-N} 2^{\frac{5}{2}k} \quad (8.29)$$

where the implicit constant is independent of k.

Hence we can split $\zeta_k^j = \zeta_{k,\text{near}}^j + \zeta_{k,\text{far}}^j$, where

$$\zeta_{k,\text{near}}^j(x) := \zeta_k^j(x) 1_{\left\{ x : 2^k |x^1| \leq L, \, 2^{\frac{k}{2}} |x'| \leq L \right\}}(x),$$

and $L > 0$ is chosen large enough (independent of k) so that, by (8.29), we have

$$\left\| \sum_{j=2}^{4} \partial_j \zeta_{k,\text{far}}^j \right\|_{L_x^{\frac{4}{3}}} \leq 2^k \varepsilon^4. \quad (8.30)$$

We denote the corresponding splitting of Z_k^j by $Z_{k,\text{near}}^j + Z_{k,\text{far}}^j$.

We are now ready to complete the proof of (8.21). The contribution of $Z_{k,\text{far}}^j \partial_j$ is acceptable, thanks to (8.22), (8.30) and the Sobolev embedding $\dot{H}_x^1 \subseteq L_x^4$. For $\sum_{j=2}^{4} Z_{k,\text{near}}^j \partial_j$, we have

$$2^{-k} \left| \sum_{j=2}^{4} Z_{k,\text{near}}^j \partial_j \psi^{(n)}(x) \right| \leq 2^{-k} \sum_{j=2}^{4} \int |\zeta_{k,\text{near}}^j(y)| |\partial_j \psi^{(n)}(x-y)| \, dy$$

$$\lesssim M w_k^{-\frac{1}{2}}(x) \| w_k^{\frac{1}{2}} \partial_j \psi^{(n)} \|_{L_x^2}$$

where, by (8.27), (8.28) and the definition of $\zeta_{k,\text{near}}^j$, M obeys the bound

$$M := \left(2^{-2k} \sum_{j=2}^{4} \int \frac{w_k(x)}{w_k(x-y)} \left| \zeta_{k,\text{near}}^j \right|^2 (y) \, dy \right)^{\frac{1}{2}}$$

$$\lesssim_L \left(2^{-2k} \sum_{j=2}^{4} \int_{\left\{ 2^k |y^1| \leq L, \, 2^{\frac{k}{2}} |y'| \leq L \right\}} |\zeta_k^j|^2 \, dy \right)^{\frac{1}{2}} \lesssim 1.$$

Note that $w_k^{-\frac{1}{2}}(x) \lesssim ((1-|x|)_+ + 2^{-k})^{1/4} + \varepsilon^4$, since we have chosen $x = (x^1, 0, 0, 0)$; this proves (8.21).

Step 3 In this step we upgrade (8.21) to the following gauge dependent estimate:

$$\limsup_{n \to \infty} 2^{-2k} |\zeta_{2^{-k}} * \mathbf{D}_j \phi^{(n)}(1, x)| \lesssim \left(2^{-\frac{3}{8}k} + \left((1-|x|)_+ + 2^{-k} \right)^{\frac{1}{4}} + \varepsilon^4 \right) E^{\frac{1}{2}}.$$

(8.31)

The idea is that (8.21) has already broken the scaling invariance, so we can easily incorporate A using the trivial bound $\|A\|_{L_t^\infty \dot{H}_x^1} \lesssim E^{1/2}$.

We begin by applying Step 2 to $\widetilde{\zeta}_{2^{-k}}$, where we recall that $\zeta = \widetilde{\zeta} * \widetilde{\zeta}$. We again introduce the shorthand $\widetilde{Z}_k(\cdot) := \widetilde{\zeta}_{2^{-k}} * (\cdot)$. By the simple pointwise inequality $|\widetilde{Z}_k \phi^{(n)}| \leq |\widetilde{Z}_k|\phi^{(n)}||$, which holds since $\widetilde{\zeta} \geq 0$, we have

$$\limsup_{n \to \infty} 2^{-k} |\widetilde{Z}_k \phi^{(n)}(1, x)| \lesssim \left(2^{-\frac{3}{8}k} + \left((1-|x|)_+ + 2^{-k} \right)^{\frac{1}{4}} + \varepsilon^4 \right) E^{\frac{1}{2}}.$$

(8.32)

Note furthermore that $Z_k = \widetilde{Z}_k^2$. For $j = 1, \ldots, 4$, we may write

$$\begin{aligned}
2^{-2k} |Z_k \mathbf{D}_j^{(n)} \phi^{(n)}(1, x)| &\leq 2^{-2k} |Z_k \partial_j \phi^{(n)}(1, x)| \\
&\quad + 2^{-2k} |Z_k (A_j^{(n)} \phi^{(n)})(1, x)| \\
&\lesssim 2^{-k} \sup_{|x-x'| \lesssim 2^{-k}} |\widetilde{Z}_k \phi^{(n)}(1, x')| \\
&\quad + 2^{-2k} |Z_k (A_j^{(n)} \phi^{(n)})(1, x)|.
\end{aligned}$$

The first term on the last line is acceptable, thanks to (8.32). To treat the second term, we insert $1 = (1 - \widetilde{Z}_{k+m}) + \widetilde{Z}_{k+m}$ in front of both $A^{(n)}$ and $\phi^{(n)}$ for some $m > 0$ to be determined. By the simple inequalities $|Z_k f(x)| \lesssim 2^{3k} \|f\|_{L_x^{4/3}}$ and $\|(1 - \widetilde{Z}_{k+m}) f\|_{L_x^2} \lesssim 2^{-k-m} \|f\|_{\dot{H}_x^1}$, as well as the Sobolev embedding $\dot{H}_x^1 \subseteq L_x^4$, we have

$$2^{-2k} |Z_k ((1 - \widetilde{Z}_{k+m}) A_j^{(n)} \cdot \phi^{(n)})(1, x)| \lesssim 2^{-m} \|A^{(n)}(1, \cdot)\|_{\dot{H}_x^1} \|\phi^{(n)}(1, \cdot)\|_{\dot{H}_x^1}$$

which can be made $\leq \varepsilon^4 E^{\frac{1}{2}}$ by choosing m large enough. Proceeding similarly, the same upper bound can be proved for $Z_k (\widetilde{Z}_{k+m} A_j^{(n)} (1 - \widetilde{Z}_{k+m}) \phi^{(n)})$.

For the remaining term, we have

$$2^{-2k} |Z_k \left(\widetilde{Z}_{k+m} A_j^{(n)} \widetilde{Z}_{k+m} \phi^{(n)} \right) (1, x)|$$

$$\lesssim 2^{-k} \|\widetilde{Z}_{k+m} A_j^{(n)}(1, \cdot)\|_{L_x^\infty} 2^{-k} \sup_{|x-x'|\lesssim 2^{-k}} |\widetilde{Z}_{k+m} \phi^{(n)}(1, x')|$$

$$\lesssim_{E,m} 2^{-k+m} \sup_{|x-x'|\lesssim 2^{-k}} |\widetilde{Z}_{k+m} \phi^{(n)}(1, x')|$$

which is acceptable in view of (8.32).

Step 4 We are ready to conclude the proof of the lemma. By (8.21) and the pointwise inequality $|\zeta_{2^{-k}} * \phi| \leq \zeta_{2^{-k}} * |\phi|$, we can achieve the desired smallness as in (8.17) of $\phi^{(n)}$ by taking K very large, γ close enough to 1 and $\varepsilon > 0$ sufficiently small. For $\mathbf{D}_t^{(n)}$, we have

$$2^{-2k} |\zeta_{2^{-k}} * \mathbf{D}_t^{(n)} \phi^{(n)}(1, x)| \leq 2^{-2k} |\zeta_{2^{-k}} * \mathbf{D}_L^{(n)} \phi^{(n)}(1, x)|$$
$$+ \sum_{j=1} 2^{-2k} |\zeta_{2^{-k}} * \mathbf{D}_j^{(n)} \phi^{(n)}(1, x)|. \quad (8.33)$$

For the first term, we begin by estimating

$$2^{-2k} |\zeta_{2^{-k}} * \mathbf{D}_L^{(n)} \phi^{(n)}(1, x)| \lesssim \left(\int_{\{y:|y-x|\lesssim 2^{-k}\}} |\mathbf{D}_L^{(n)} \phi^{(n)}(1, y)|^2 \, \mathrm{d}y \right)^{1/2}.$$

Then by (8.20), the right-hand side is bounded by

$$\left(\left((1 - |x|)_+ + 2^{-k} \right)^{\frac{1}{4}} + \varepsilon^4 \right) E^{1/2}$$

provided that $\varepsilon_n^{1/2} \leq \frac{1}{10} \varepsilon^8$. Using (8.31) to estimate the second term in (8.33), (8.17) now follows after adjusting K, γ and ε if necessary. $\qquad\square$

8.3 Nontrivial energy in a time-like region

An important consequence of Lemma 8.7 is that there is a uniform lower bound for $\phi^{(n)}$ in a time-like region at $t = 1$.

Lemma 8.9 *Let $(A^{(n)}, \phi^{(n)})$ be a sequence of admissible $C_t \mathcal{H}^1$ solutions to* (MKG) *satisfying the conclusions of Lemma 8.4. Let $K(E) > 0$ and $\gamma(E) \in (0, 1)$ be as in Lemma 8.7 and Remark 8.8. Assume that either* (1) $k_n \leq K(E)$ *or* (2) $k_n > K(E)$ *and* $|x_n| \leq \gamma(E)$. *Then there exist $E_1 = E_1(E) > 0$ and $\gamma_1 = \gamma_1(E) \in (0, 1)$ such that if $\varepsilon > 0$ is sufficiently small depending on E, then*

$$\int_{S_1^{1-\gamma_1}} \sum_{\mu=0}^{4} |\mathbf{D}_\mu^{(n)} \phi^{(n)}|^2 + \frac{1}{r^2} |\phi^{(n)}|^2 \, \mathrm{d}x \geq E_1(E). \tag{8.34}$$

Proof Since the whole proof will take place on $\{t = 1\}$, we will ignore the difference between $\{t = 1\}$ and $\mathbb{R}^4$. In this case, note that $S_1^{1-\gamma} = \{1\} \times B_\gamma(0)$ for any $\gamma \in (0, 1)$. Furthermore, as the argument is the same for each n, we will henceforth suppress n for simplicity.

There are two scenarios to consider:

A. **Nontrivial kinetic energy.** $2^{-2k} |\zeta_{2^{-k}} * \mathbf{D}_t \phi(x)| \geq \frac{1}{2} e(E)$, or
B. **Nontrivial potential energy.** $2^{-k} |\zeta_{2^{-k}} * \phi(x)| \geq \frac{1}{2} e(E)$.

We first treat Scenario A. By Cauchy-Schwarz,

$$\frac{1}{2} e \leq \int 2^{-2k} \zeta_{2^{-k}}(y) |\mathbf{D}_t \phi(x - y)| \, \mathrm{d}y \lesssim \left(\int_{B_{2^{-k}}(x)} |\mathbf{D}_t \phi|^2 \, \mathrm{d}y \right)^{1/2},$$

where we also used $\operatorname{supp} \zeta \subseteq B_1(0)$. Hence in Case 2, (8.34) immediately follows by taking $\gamma_1 \geq \gamma + 2^{-k}$ so that $\{1\} \times B_{2^{-k}}(x) \subseteq S_1^{1-\gamma_1}$. Note that we may still ensure that $\gamma_1 < 1$ thanks to (8.18).

Now assume that Case 1 holds, i.e., $k \leq K$. Splitting the convolution integral into $\int_{S_1^{1-\gamma_1}} + \int_{S_1 \setminus S_1^{1-\gamma_1}} + \int_{\mathbb{R}^4 \setminus S_1}$, applying Cauchy-Schwarz and using (8.9), (8.10), we have

$$\mathbf{e} \lesssim \left(\int_{S_1^{1-\gamma_1}} |\mathbf{D}_t \phi|^2 \, \mathrm{d}y \right)^{1/2} + c_0(\gamma_1) E^{1/2} + \varepsilon^8 E^{1/2},$$

where

$$c_0(\gamma_1) := \left(\int_{S_1 \setminus S_1^{1-\gamma_1}} |\zeta(2^{-k} y)|^2 2^{-4k} \, \mathrm{d}y \right)^{1/2}$$
$$\lesssim 2^{-2k} | \left(S_1 \setminus S_1^{1-\gamma_1} \right) \cap B_{2^{-k}}(x) |^{1/2}.$$

By elementary geometry and the assumption $k \leq K$, it follows that the last term is bounded by $\lesssim (1 - \gamma_1)^{1/2} 2^{-K/2}$ uniformly in x. Taking γ_1 sufficiently close to 1, the desired conclusion follows.

We now consider Scenario B. We repeat the above argument with $\mathbf{D}_t \phi$ replaced by ϕ, while putting $\zeta_{2^{-k}}$ [resp. ϕ] in $L^{4/3}$ [resp. L^4] instead of L^2 [resp. L^2]. Then in Case 1,

$$\mathbf{e} \lesssim \left(\int_{B_{2^{-k}}(x)} |\phi|^4 \, dy \right)^{1/4}, \tag{8.35}$$

whereas in Case 2,

$$\mathbf{e} \lesssim \left(\int_{S_1^{1-\gamma_1}} |\phi|^4 \, dy \right)^{\frac{1}{4}} + c_1(\gamma_1) \|\phi\|_{L_x^4(\mathbb{R}^4)} + \|\phi\|_{L_x^4(\mathbb{R}^4 \setminus S_1)}, \tag{8.36}$$

with $c_1(\gamma_1) \lesssim (1 - \gamma_1)^{3/4} 2^{-3K/4}$. The desired conclusion then follows from (8.9), (8.10), the diamagnetic inequality (Lemma 8.2) and the localized Sobolev inequalities

$$\|f\|_{L_x^4(B_r(0))} \lesssim \left(\sum_{j=1}^4 \|\partial_j f\|_{L_x^2(B_r(0))}^2 \right)^{1/2} + \left\| \frac{1}{|x|} f \right\|_{L^2(B_r(0))},$$

$$\|f\|_{L_x^4(\mathbb{R}^4 \setminus B_r(0))} \lesssim \left(\sum_{j=1}^4 \|\partial_j f\|_{L_x^2(\mathbb{R}^4 \setminus B_r(0))}^2 \right)^{1/2},$$

which hold with a uniform constant for any $r > 0$.

To prove the preceding two inequalities, it suffices to consider the case $r = 1$ by scaling invariance. The first inequality is an immediate consequence of the standard inequality $\|f\|_{L^4(B_1(0))} \lesssim \|f\|_{\dot{H}^1(B_1(0))} + \|f\|_{L^2(B_1(0))}$. To prove the second inequality, we extend f to $\mathbb{R}^4$. Using the standard extension operator from f on $B_2(0) \setminus B_1(0)$, the global extension $\widetilde{f}$ on $\mathbb{R}^4$ can be chosen so that $\widetilde{f} = f$ on $\mathbb{R}^4 \setminus B_1(0)$ and

$$\|\widetilde{f}\|_{H^1(B_1(0))} \lesssim \|f\|_{H^1(B_2(0) \setminus B_1(0))} \lesssim \|f\|_{\dot{H}^1(\mathbb{R}^4 \setminus B_1(0))} + \|\frac{1}{|x|} f\|_{L^2(\mathbb{R}^4 \setminus B_1(0))}.$$

Using localized Hardy's inequality in Lemma 5.8, the second term on the right-hand side may be bounded by the first term. The desired localized Sobolev inequality now follows from the usual Sobolev embedding $\dot{H}^1 \subseteq L^4$. $\qquad\square$

Combining Lemmas 8.7 and 8.9, it follows that any sequence $(A^{(n)}, \phi^{(n)})$ of admissible $C_t \mathcal{H}^1$ solutions to (MKG) constructed by Lemma 8.4 obeys the uniform lower bound (8.34) in a time-like region $S_1^{1-\gamma_1}$ for some $\gamma_1 = \gamma_1(E) \in (0, 1)$. The uniform lower bound in a time-like region can be propagated towards $t = 0$ using the localized monotonicity formula in Proposition 5.5.

Lemma 8.10 *Let $(A^{(n)}, \phi^{(n)})$ be a sequence of admissible $C_t \mathcal{H}^1$ solutions to* (MKG) *satisfying the conclusions of Lemma 8.4. Assume furthermore that*

each $(A^{(n)}, \phi^{(n)})$ obeys (8.34). *Then there exist $E_2 = E_2(E) > 0$ and $\gamma_2 = \gamma_2(E) \in (0, 1)$ such that*

$$\int_{S_t^{(1-\gamma_2)t}} {}^{(X_0)} P_T[A^{(n)}, \phi^{(n)}] \, dx \geq E_2(E) \quad \text{for every } t \in \left[\varepsilon_n^{\frac{1}{2}}, \varepsilon_n^{\frac{1}{4}}\right]. \quad (8.37)$$

Proof Fix n and $t_0 \in [\varepsilon_n^{1/2}, \varepsilon_n^{1/4}]$. Applying Proposition 5.5 with $\varepsilon = \varepsilon_n$, $\delta_0 = (1 - \gamma_2)t_0$ and $\delta_1 = M\delta_0$, where $\gamma_2 \in (0, 1)$ and $M > 1$ will be chosen below, we obtain

$$\int_{S_1^{M(1-\gamma_2)t_0}} {}^{(X_0)} P_T[A, \phi] \, dx \leq \int_{S_{t_0}^{(1-\gamma_2)t_0}} {}^{(X_0)} P_T[A, \phi] \, dx$$
$$+ C\left((M(1-\gamma_2))^{\frac{1}{2}} + |\log M|^{-1}\right)E. \quad (8.38)$$

On the other hand, by Lemma 5.10 (in particular, the expression for ${}^{(X_0)} P_T = \frac{1}{2}({}^{(X_0)} P_L + {}^{(X_0)} P_{\underline{L}})$ and (8.34), we have

$$E_1 \lesssim (1 - \gamma_1)^{-\frac{1}{2}} \int_{S_1^{1-\gamma_1}} {}^{(X_0)} P_T[A, \phi] \, dx.$$

Hence choosing M sufficiently large and γ_2 close enough to 1 to make the last term in (8.38) small, (8.37) follows with $E_2 = cE_1(1 - \gamma_1)^{\frac{1}{2}}$ for some $c > 0$.
$\square$

8.4 Final rescaling

So far, under the assumption that Theorem 1.3 fails, we have shown the existence of a sequence of solutions $(A^{(n)}, \phi^{(n)})$ that satisfies the conclusions of Lemma 8.4 and a uniform lower bound (8.37) in a time-like region. By Proposition 5.4, the sequence moreover obeys the uniform space-time bound

$$\iint_{C_{[\varepsilon_n, 1]}} \frac{1}{\rho_{\varepsilon_n}} |\iota_{X_{\varepsilon_n}} F^{(n)}|^2 + \frac{1}{\rho_{\varepsilon_n}} \left|\left(\mathbf{D}_{X_{\varepsilon_n}}^{(n)} + \frac{1}{\rho_{\varepsilon_n}}\right)\phi^{(n)}\right|^2 \, dt dx \lesssim E. \quad (8.39)$$

Our next goal is to upgrade (8.39) to asymptotic self-similarity by a rescaling argument.

Lemma 8.11 *Suppose that Theorem 1.3 fails. Then there exists a sequence of admissible $C_t \mathcal{H}^1$ solutions $(A^{(n)}, \phi^{(n)})$ on $[1, T_n] \times \mathbb{R}^4$ with $T_n \to \infty$ satisfying the following properties:*

(1) *Bounded energy in the cone*

$$\mathcal{E}_{S_t}\left[A^{(n)}, \phi^{(n)}\right] \le E, \qquad \text{for every } t \in [1, T_n], \qquad (8.40)$$

(2) *Small energy outside the cone*

$$\mathcal{E}_{\{t\} \times \mathbb{R}^4 \setminus S_t}\left[A^{(n)}, \phi^{(n)}\right] \le \frac{1}{100} E \qquad \text{for every } t \in [1, T_n], \quad (8.41)$$

(3) *Nontrivial energy in a time-like region*

$$\int_{S_t^{(1-\gamma_2)t}} {}^{(X_0)} P_T\left[A^{(n)}, \phi^{(n)}\right] \mathrm{d}x \ge E_2 \quad \text{for every } t \in [1, T_n], \quad (8.42)$$

(4) *Asymptotic self-similarity*

$$\iint_K |\iota_{X_0} F^{(n)}|^2 + \left|\left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho}\right)\phi^{(n)}\right|^2 \mathrm{d}t\mathrm{d}x \to 0 \quad \text{as } n \to \infty \quad (8.43)$$

for every compact subset K of the interior of $C_{[1,\infty)}$.

Proof Let $(A^{(n)}, \phi^{(n)})$ be a sequence of solutions satisfying the conclusions of Lemmas 8.4 and 8.10. Consider the time interval $[\varepsilon_n^{1/2}, \varepsilon_n^{1/4}]$, on which (8.37) applies. Given $T_n > 1$, we partition ε_n in to dyadic intervals of the form $I_n^j = [T_n^j \varepsilon_n^{1/2}, T_n^{j+1} \varepsilon_n^{1/2}]$; there are roughly $|\log \varepsilon_n|/\log T_n$ many such intervals. We choose T_n so that $\log T_n \sim |\log \varepsilon_n|^{1/2}$. Observe that $T_n \to \infty$. Also, by the pigeonhole principle applied to (8.39), there exists $j(n)$ such that

$$\iint_{C_{I_n^{j(n)}}} \frac{1}{\rho_{\varepsilon_n}} |\iota_{X_{\varepsilon_n}} F^{(n)}|^2 + \frac{1}{\rho_{\varepsilon_n}} \left|\left(\mathbf{D}_{X_{\varepsilon_n}}^{(n)} + \frac{1}{\rho_{\varepsilon_n}}\right)\phi^{(n)}\right|^2 \mathrm{d}t\mathrm{d}x$$

$$\lesssim \frac{\log T_n}{|\log \varepsilon_n|} E \sim \frac{1}{|\log \varepsilon_n|^{1/2}} E, \qquad (8.44)$$

which decays to 0 as $n \to \infty$.

We now rescale $C_{I_n^{j(n)}}$ to $C_{[1,T_n]}$; abusing the notation a bit (but conforming to the statement of the lemma), we denote the rescaled solutions again by $(A^{(n)}, \phi^{(n)})$. From (8.9) and (8.10) with $\varepsilon^8 \le \frac{1}{100}$, (8.40) and (8.41) follow. Also, (8.42) is a consequence of (8.37). Furthermore, (8.44) implies

$$\iint_{C_{[1,T_n]}} \frac{1}{\rho_{\varepsilon_n'}} |\iota_{X_{\varepsilon_n'}} F^{(n)}|^2 + \frac{1}{\rho_{\varepsilon_n'}} \left| \left(\mathbf{D}_{X_{\varepsilon_n'}}^{(n)} + \frac{1}{\rho_{\varepsilon_n'}} \right) \phi^{(n)} \right|^2 \, dt dx \to 0 \quad \text{as } n \to \infty$$

(8.45)

where $\varepsilon_n' := (T_n^{j(n)} \varepsilon_n^{1/2})^{-1} \varepsilon_n$ obeys $\varepsilon_n' \le \varepsilon_n^{1/2} \to 0$. For any compact subset K of the interior of $C_{[1,\infty)}$, which is in particular situated away from the boundary $\partial C_{[1,\infty)}$, we claim that

$$\iint_K \left(\frac{1}{\rho_{\varepsilon_n'}} |\iota_{X_{\varepsilon_n'}} F^{(n)}|^2 - \frac{1}{\rho} |\iota_{X_0} F^{(n)}|^2 \right)$$
$$+ \left(\frac{1}{\rho_{\varepsilon_n'}} \left| \left(\mathbf{D}_{X_{\varepsilon_n'}}^{(n)} - \frac{1}{\rho_{\varepsilon_n'}} \right) \phi^{(n)} \right|^2 - \frac{1}{\rho} \left| \left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 \right) dt dx \to 0$$

Indeed, in the coordinates $(x^0 = t, x^1, \ldots, x^4)$, the left-hand side can be written in the form

$$\iint_K d_1^{(n)\mu\nu} |F_{\mu\nu}^{(n)}|^2 + d_2^{(n)\mu} |\mathbf{D}_\mu^{(n)} \phi^{(n)}|^2 + d_3^{(n)} |\phi^{(n)}|^2 \, dt dx$$

where $d_1^{(n)\mu\nu}(t, x)$, $d_2^{(n)\mu}(t, x)$ and $d_3^{(n)}$ are continuous functions which tend to 0 pointwisely (hence uniformly) on K, whereas $|F_{\mu\nu}|$, $|\mathbf{D}_\mu^{(n)} \phi^{(n)}|$ and $|\phi^{(n)}|$ are uniformly in $L^2(K)$ by (8.40), (8.41) and Hardy's inequality. By Hölder's inequality, the claim follows. Then combining the claim with (8.45), we arrive at the desired asymptotic self-similarity (8.43). □

8.5 Concentration scales

Let $(A^{(n)}, \phi^{(n)})$ be a sequence of solutions given by Lemma 8.11. We now present a combinatorial result that establishes the following dichotomy: Either there is a uniform non-concentration of energy, or we can identify a sequence of points and decreasing scales at which energy concentrates.

To state the result, we need few definitions. For each $j = 1, 2, \ldots$ we define

$$C_j := \left\{ (t, x) \in C_{[1,\infty)}^1 : 2^j \le t < 2^{j+1} \right\},$$
$$\widetilde{C}_j := \left\{ (t, x) \in C_{[1/2,\infty)}^{1/2} : 2^j \le t < 2^{j+1} \right\}.$$

In words, C_j [resp. $\widetilde{C}_j$] is the set of points in the truncated cone $C_{[2^j, 2^{j+1})}$ at distance ≥ 1 [resp. $\geq 1/2$] from the lateral boundary. For each $j \geq 1$, we have the following lemma.

Lemma 8.12 *Let $(A^{(n)}, \phi^{(n)})$ be a sequence of admissible $C_t \mathcal{H}^1$ solutions on $[1, T_n] \times \mathbb{R}^4$ with $T_n \to \infty$ satisfying (8.40)–(8.43) for some $E > 0$. Let ϵ_0 be as in Proposition 6.1. Then for each $j = 1, 2, \cdots$, after passing to a subsequence, one of the following alternatives holds:*

(1) **Concentration of energy**. *There exist points $(t_n, x_n) \in \widetilde{C}_j$, scales $r_n \to 0$ and $0 < r = r(j) < 1/4$ such that the following bounds hold:*

$$\mathcal{E}_{\{t_n\} \times B_{r_n}(x_n)} \left[A^{(n)}, \phi^{(n)} \right] = \frac{1}{C_0^2} \epsilon_0^2, \tag{8.46}$$

$$\sup_{x \in B_r(x_n)} \mathcal{E}_{\{t_n\} \times B_{r_n}(x)} \left[A^{(n)}, \phi^{(n)} \right] \leq \frac{1}{C_0^2} \epsilon_0^2, \tag{8.47}$$

$$\frac{1}{4r_n} \int_{t_n - 2r_n}^{t_n + 2r_n} \int_{B_r(x_n)} |\iota_{X_0} F^{(n)}|^2$$
$$+ \left| \left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 \, dt dx \to 0 \quad as \ n \to \infty. \tag{8.48}$$

(2) **Uniform non-concentration of energy**. *There exists $0 < r = r(j) < 1/4$ such that the following bounds hold:*

$$\int_{S_t^{(1-\gamma_2)t}} {}^{(X_0)} P_T \left[A^{(n)}, \phi^{(n)} \right] \, dx \geq E_2 \quad for \ t \in [2^j, 2^{j+1}), \tag{8.49}$$

$$\sup_{(t,x) \in C_j} \mathcal{E}_{\{t\} \times B_r(x)} \left[A^{(n)}, \phi^{(n)} \right] \leq \frac{1}{C_0^2} \epsilon_0^2, \tag{8.50}$$

$$\iint_{\widetilde{C}_j} |\iota_{X_0} F^{(n)}|^2 + \left| \left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 \, dt dx \to 0 \quad as \ n \to \infty. \tag{8.51}$$

Here $C_0 > 0$ is a universal constant much larger than the implicit constants in Lemma 4.5.

Proof This lemma is essentially [33, Lemma 6.3]; for completeness we give a self-contained alternative proof, which relies on the use of the Hardy-Littlewood maximal function theorem to establish (8.48).

   Step 1 Fix $j \in \{1, 2, \ldots\}$. We begin by identifying a 'low energy barrier' around C_j inside $\widetilde{C}_j$. Let $N > 0$ be a large integer to be determined later. We first partition the time interval $[2^j, 2^{j+1})$ into smaller intervals I_k, where

$$I_k := \left[2^j + \frac{k-1}{10N}, 2^j + \frac{k}{10N} \right) \quad k = 1, \dots, 10N2^j.$$

Accordingly, define $C_j^k := C_j \cap (I_k \times \mathbb{R}^4)$ and $\widetilde{C}_j^k := \widetilde{C}_j \cap (I_k \times \mathbb{R}^4)$. Next, we partition $\widetilde{C}_j^k \setminus C_j^k$ into $\cup_{\ell=1}^N \widetilde{C}_j^{k,\ell}$, where

$$\widetilde{C}_j^{k,\ell} = \left\{ (t,x) \in \widetilde{C}_j^k : \frac{1}{2} + \frac{\ell-1}{2N} \le t - |x| < \frac{1}{2} + \frac{\ell}{2N} \right\}, \quad \ell = 1, \dots, N.$$

For each n and k, we claim that there exists $1 \le \ell(n,k) \le N$ such that

$$\sup_{t \in I_k} \mathcal{E}_{S_t \cap \widetilde{C}_j^{k,\ell(n,k)}} \left[A^{(n)}, \phi^{(n)} \right] \le \frac{3}{N} E. \tag{8.52}$$

Indeed, for each k consider the left endpoint $t_k := 2^j + (k-1)/(10N)$. The set $S_{t_k} \cap (\widetilde{C}_j^k \setminus C_j^k)$ is partitioned into N annuli of the form $S_{t_k} \cap \widetilde{C}_j^{k,\ell}$. By the pigeonhole principle and the energy bound (8.40), there exists $1 \le \ell(n,k) \le N-2$ such that

$$\sum_{\ell=\ell(n,k)}^{\ell(n,k)+2} \mathcal{E}_{S_{t_k} \cap \widetilde{C}_j^{k,\ell}} \left[A^{(n)}, \phi^{(n)} \right] \le \frac{3}{N} E.$$

As $\widetilde{C}_j^{k,\ell(n,k)}$ lies in the domain of dependence of $\cup_{\ell=\ell(n,k)}^{\ell(n,k)+2} S_{t_k} \cap \widetilde{C}_j^{k,\ell}$, (8.52) now follows by the local conservation of energy.

We choose N large enough so that

$$\frac{3}{N} E \le \frac{1}{C_0^2} \epsilon_0^2.$$

Hence, by (8.52), $\widetilde{C}_j^{k,\ell(n,k)}$ serves as a 'low energy barrier' that separates the behavior of the solution in the interior $\widetilde{C}_j^{k,<\ell(n,k)} := (\cup_{\ell=1}^{\ell(n,k)-1} \widetilde{C}_j^{k,\ell}) \cup C_j^k$ from the outside. Fix $r_0 = \frac{1}{4N}$, so that $0 < r_0 < 1/4$ and

$$(t,x) \in \widetilde{C}_j^{k,<\ell(n,k)} \Rightarrow \{t\} \times B_{4r_0}(x) \subseteq \widetilde{C}_j^{k,<\ell(n,k)} \cup \widetilde{C}_j^{k,\ell(n,k)} \subseteq C_{[1/2,\infty)}^{1/2}. \tag{8.53}$$

Step 2 For each n and k, define $f_{n,k} : [0, r_0] \times I_k \to [0, \infty)$ by

$$f_{n,k}(r,t) := \sup \left\{ \mathcal{E}_{\{t\} \times B_r(x)} [A^{(n)}, \phi^{(n)}] : (t,x) \in \widetilde{C}_j^{k,<\ell(n,k)} \right\}.$$

We then define the lowest energy concentration scale $r_{n,k}(t)$ as

$$r_{n,k}(t) := \begin{cases} \inf\left\{r \in [0, r_0] : f_{n,k}(t, r) \geq \frac{1}{C_0^2}\epsilon_0^2\right\} & \text{if } f_{n,k}(t, r_0) \geq \frac{1}{C_0^2}\epsilon_0^2, \\ r_0 & \text{otherwise.} \end{cases}$$

$$(8.54)$$

We claim that each $r_{n,k}$ is Lipschitz continuous with a universal constant $c_L > 0$, i.e.,

$$|r_{n,k}(t_1) - r_{n,k}(t_0)| \leq c_L|t_1 - t_0| \quad \text{for } t_0, t_1 \in I_k.$$

The key idea is to use the finite speed of propagation, or equivalently, local conservation of energy. Let $t_0, t_1 \in I_k$ with $t_0 < t_1$. Consider first the case when $r_{n,k}(t_0) \geq r_{n,k}(t_1)$. For convenience, we introduce the shorthand $\bar{r} := r_{n,k}(t_1)$. When $\bar{r} = r_0$, then necessarily $r_{n,k}(t_0) = r_0$ and (8.5) holds trivially. If $\bar{r} < r_0$, then there exists $x_1 \in \mathbb{R}^4$ such that $(t_1, x_1) \in \widetilde{C}_j^{k, < \ell(n,k)}$ and $\mathcal{E}_{\{t_1\} \times B_{\bar{r}}}[A^{(n)}, \phi^{(n)}] = \frac{1}{C_0^2}\epsilon_0^2$. By local conservation of energy, it follows that

$$\mathcal{E}_{\{t_0\} \times B_{\bar{r} + (t_1 - t_0)}(x_1)}\left[A^{(n)}, \phi^{(n)}\right] \geq \mathcal{E}_{\{t_1\} \times B_{\bar{r}}}\left[A^{(n)}, \phi^{(n)}\right] = \frac{1}{C_0^2}\epsilon_0^2.$$

If $(t_0, x_1) \in \widetilde{C}_j^{k, < \ell(n,k)}$, then $r_{n,k}(t_0) \leq \bar{r} + (t_1 - t_0)$. If $(t_0, x_1) \notin \widetilde{C}_j^{k, < \ell(n,k)}$, then by elementary geometry there exists $(t_0, x_0) \in \widetilde{C}_j^{k, < \ell(n,k)}$ such that $|x_1 - x_0| < t_1 - t_0$. Hence the energy of $(A^{(n)}, \phi^{(n)})$ on $\{t_0\} \times B_{\bar{r} + 2(t_1 - t_0)}(x_0)$ is bounded from below by $\frac{1}{C_0^2}\epsilon_0^2$, which implies $r_{n,k}(t_0) \leq \bar{r} + 2(t_1 - t_0)$ in general. Treating the other case $r_{n,k}(t_0) > r_{n,k}(t_1)$ in a similar way, it follows that (8.5) holds with $c_L = 2$.

We now proceed to the proof of of the lemma. We first treat the case when there exists a common lower bound $0 < r(j) \leq r_0$ of $r_{n,k}$, i.e., $r_{n,k}(t) \geq r(j)$ for all n, k and $t \in I_k$. Unraveling the definition of $r_{n,k}$, we see that (8.50) holds. Moreover, (8.49) and (8.51) follow directly from (8.42) and (8.43), respectively. Thus we conclude that the second scenario (uniform non-concentration of energy) holds.

To complete the proof, it only remains to consider the alternative case and show that the first scenario (concentration of energy) holds. After passing to a subsequence, we may assume that there exists $k \in \{1, \ldots, 10N2^j\}$ such that

$$\liminf_{n \to \infty} \inf_{I_k} r_{n,k} = 0. \quad (8.55)$$

Then we claim that there exist (t_n, x_n) and r_n such that (8.46)–(8.48) hold with $r(j) = r_0$, up to passing to a subsequence.

Define

$$\alpha_n^2 := \int_{2^{j-1}}^{2^{j+2}} \beta_n^2(t) \, dt,$$

$$\beta_n^2(t) := \int_{S_t \cap C_{[1/2,\infty)}^{1/2}} |\iota_{X_0} F^{(n)}|^2 + \left| \left(\mathbf{D}_{X_0}^{(n)} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 \, dx.$$

Note that $\alpha_n^2 \to 0$ by (8.43). By the Hardy-Littlewood maximal function theorem, for every $\alpha > 0$ we have

$$\left| \left\{ t \in [2^{j-1}, 2^{j+1}) : M\left[\beta_n^2\right](t) > \alpha \right\} \right| \lesssim \frac{1}{\alpha} \alpha_n^2, \qquad (8.56)$$

where $M[\beta_n](t)$ is the Hardy-Littlewood maximal function on $[2^{j-1}, 2^{j+2})$, given by

$$M[\beta_n](t) := \sup_{a>0} \frac{1}{2a} \int_{(t-a,t+a) \cap [2^{j-1}, 2^{j+2})} \beta_n^2(t') \, dt'.$$

Roughly speaking (8.56) says that the desired conclusion (8.48) holds for 'most of' $t \in I_k$. This fact, combined with the flexibility of the choice of t_n such that $\lim_{n \to \infty} r_{n,k}(t_n) = 0$, will lead to the desired conclusions (8.46)–(8.48).

More precisely, define the intervals $J_n, K_n \subseteq I_k$ by

$$J_n := \left\{ t \in I_k : M[\beta_n^2](t) \leq \alpha_n \right\}, \quad K_n := (\bar{t}_n - \alpha_n^{1/2}, \bar{t}_n + \alpha_n^{1/2}) \cap I_k,$$

where $\bar{t}_n \in I_k$ is a minimum of $r_{n,k}$, i.e., $r_{n,k}(\bar{t}_n) = \inf_{I_k} r_{n,k}$. By the uniform Lipschitz continuity of $r_{n,k}$ and the fact that $\alpha_n^2 \to 0$ as $n \to \infty$, we have

$$\sup_{t \in K_n} r_{n,k}(t) \to 0 \quad \text{as } n \to \infty.$$

Note that $|I_k \setminus J_n| \lesssim \alpha_n$ by (8.56) with $\alpha = \alpha_n$, whereas $|K_n| = 2\alpha_n^{1/2}$. Using again the fact that $\alpha_n^2 \to 0$ as $n \to \infty$ and passing to a subsequence, it follows that $J_n \cap K_n \neq \emptyset$ for all n. Choosing t_n so that $t_n \in J_n \cap K_n$ and $r_n := r_{n,k}(t_n)$, we have

$$\sup_{a>0} \frac{1}{2a} \int_{t_n-a}^{t_n+a} \beta_n^2(t) \, dt \to 0, \quad r_n = r_{n,k}(t_n) \to 0 \quad \text{as } n \to \infty.$$

With the choice $r(j) = r_0$, (8.48) follows from (8.53) and the previous state-
ment. Passing to a subsequence if necessary, we may assume that $r_{n,k}(t_n) < r_0$;
then there exists $(t_n, x_n) \in \widetilde{C}_j^{k, <\ell(n,k)}$ such that (8.46) holds for all n as well.
Finally, thanks to the low energy barrier (8.53) and the definition of $r_{n,k}$, (8.47)
follows with $r(j) = r_0$.   □

8.6 Compactness/rigidity argument

We are now ready to complete the proof of Theorem 1.3, by using the tools
developed in Sects. 6 and 7.

Completion of proof of Theorem 1.3 Let $(A^{(n)}, \phi^{(n)})$ be a sequence of admis-
sible $C_t \mathcal{H}^1$ solutions on $[1, T_n] \times \mathbb{R}^4$ given by Lemma 8.11. We consider two
cases according to Lemma 8.12, and show that both lead to contradictions.

Case 1 Suppose that there exists $j \in \{1, 2, \ldots\}$ such that the first scenario
(concentration of energy) in Lemma 8.12 holds. We need to set things up so
that we can use Proposition 6.1, and for that we also need local control of the
L^2 norm of ϕ. This is achieved via the improved form of Hardy's inequality
in Lemma 4.5. From (8.47), we obtain

$$(\sigma^{-1} r_n)^{-2} \|\phi^{(n)}(t_n)\|^2_{L^2_x \left(B_{8\sigma^{-1} r_n}\right)} \leq c\epsilon_0^2 + C\sigma^{-2} E,$$

with a universal constant $c \ll 1$ and a parameter $\sigma \geq 2$ to be specified. To
eliminate the second term, we choose σ so that

$$C\sigma^{-2} E = c\epsilon_0^2.$$

Thus we have insured that the hypothesis of Proposition 6.1 are satisfied with
respect to the rescaled ball $B_{\sigma^{-1} r_n}(x)$ with x as in (8.47), i.e.,

$$\mathcal{E}_{\{t_n\} \times B_{8\sigma^{-1} r_n}(x)} \left[A^{(n)}, \phi^{(n)}\right] + \left(\sigma^{-1} r_n\right)^{-2} \|\phi^{(n)}(t_n)\|^2_{L^2_x(B_{8\sigma^{-1} r_n}(x))} \leq \epsilon_0^2$$

$$(8.57)$$

for every $x \in B_{r(j)}(x_n)$.

As $\widetilde{C}_j$ is pre-compact, we may assume that (t_n, x_n) has a limit (t_0, x_0) in
the closure of $\widetilde{C}_j$ after passing to a subsequence. Consider the sequence

$$(\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)})(t, x) := \sigma^{-1} r_n(A^{(n)}, \phi^{(n)}) \left(\sigma^{-1} r_n t + t_n, \sigma^{-1} r_n x + x_n\right).$$

By (8.46), there is always a nontrivial amount of energy at the origin, i.e.,

$$\mathcal{E}_{\{0\} \times B_\sigma(0)} \left[\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)} \right] = \frac{1}{C_0^2} \epsilon_0^2. \tag{8.58}$$

Fix any $x \in \mathbb{R}^4$. As $r_n \to 0$, observe that the point $\sigma^{-1} r_n x + x_n$ belongs to $B_{r(j)}(x_n)$ for sufficiently large n. Hence, by (8.57), we have

$$\mathcal{E}_{\{0\} \times B_8(x)} \left[\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)} \right] + \|\widetilde{\phi}^{(n)}(0)\|^2_{L^2_x(B_8(x))} \leq \epsilon_0^2 \quad \text{for sufficiently large } n. \tag{8.59}$$

Finally, by (8.48), the convergence $(t_n, x_n) \to (t_0, x_0)$ and smoothness of X_0, it follows that

$$\iint_{(-2,2) \times B_2(x)} |\iota_Y \widetilde{F}^{(n)}|^2 + |\mathbf{D}_Y^{(n)} \widetilde{\phi}^{(n)}|^2 \, dt dx \to 0 \quad \text{as } n \to \infty. \tag{8.60}$$

where $Y = X_0(t_0, x_0)$ is a constant time-like vector field. Note that the contribution of the term $\frac{1}{\rho} \phi^{(n)}$ drops out by scaling.

As a consequence, for each $x \in \mathbb{R}^4$ we can apply Proposition 6.1 to obtain a weak solution $(A_{[x]}, \phi_{[x]}) \in \mathcal{X}^w((-1, 1) \times B_1(x))$ to (MKG) such that

$$\iota_Y F_{[x]} = 0, \quad \mathbf{D}_{[x]Y} \phi_{[x]} = 0,$$

and $(\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)})$ converges to $(A_{[x]}, \phi_{[x]})$ up to gauge transformations on $(-1, 1) \times B_1(x)$ as in (6.3), (6.4). By Lemma 6.15, the weak solutions $(A_{[x]}, \phi_{[x]})$ form weak compatible pairs (as in Definition 6.13) on the open cover $\{(-1, 1) \times B_1(x)\}_{x \in (1/2)\mathbb{Z}^4}$ of $(-1, 1) \times \mathbb{R}^4$. Furthermore, by Proposition 7.3, there exists an equivalent set of smooth compatible pairs $(A_{[\alpha]}, \phi_{[\alpha]})$ on some refined open cover $\mathcal{Q} = \{Q_\alpha\}$ of $(-1, 1) \times \mathbb{R}^4$.

Let (A, ϕ) be a global smooth pair on $(-1, 1) \times \mathbb{R}^4$ equivalent to $(A_{[\alpha]}, \phi_{[\alpha]})$. We then extend (A, ϕ) to $\mathbb{R}^{1+4}$ as a smooth solution to (MKG) satisfying $\iota_Y F = 0$ and $\mathbf{D}_Y \phi = 0$ by pulling back along the flow of Y.

Note that (A, ϕ) has finite energy (in fact, bounded by $\leq E$), as we have

$$^{(T)} J_T \left[\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)} \right] \to \, ^{(T)} J_T[A, \phi] \quad \text{locally in } L^1_{t,x} \text{ on } (-1, 1) \times \mathbb{R}^4 \tag{8.61}$$

by (6.4) and the gauge invariance of the energy density $^{(T)} J_T$. After applying a suitable Lorentz transform, we may furthermore assume that $Y = T$. By Proposition 7.1, it follows that $\mathcal{E}[A, \phi] = 0$, but this contradicts (8.58) and (8.61).

Case 2 Suppose that for every $j \in \{1, 2, \ldots\}$ the second scenario (uniform non-concentration of energy) in Lemma 8.12 holds. The goal in this case is to apply results in Sect. 6 to extract a smooth nontrivial self-similar solution with finite energy, which would contradict Proposition 7.2.

Fix $j \in \{1, 2, \ldots\}$ and consider any point $(t, x) \in C^2_{[2,\infty)} \cap C_j$. By (8.50) and Lemma 4.5, where $\sigma \geq 2$ is chosen as in Case 1, we obtain the following analogue of (8.57):

$$\mathcal{E}_{\{t\} \times B_{8\sigma^{-1}r(j)}(x)} \left[A^{(n)}, \phi^{(n)} \right] + (\sigma^{-1}r(j))^{-2} \|\phi^{(n)}\|^2_{L^2_x(B_{8\sigma^{-1}r(j)}(x))} \leq \epsilon_0^2. \tag{8.62}$$

Since (t, x) belongs to the smaller cone $C^2_{[2,\infty)}$, we have

$$\widetilde{K}^j_{[t,x]} := \left(t - 2\sigma^{-1}r(j), t + 2\sigma^{-1}r(j) \right) \times B_{8\sigma^{-1}r(j)}(x)$$
$$\subseteq \widetilde{C}_{j-1} \cup \widetilde{C}_j \cup \widetilde{C}_{j+1}.$$

Therefore, by (8.51) we have

$$\iint_{\widetilde{K}^j_{[t,x]}} |\iota_{X_0} F^{(n)}|^2 + \left| \left(\mathbf{D}_{X_0} + \frac{1}{\rho} \right) \phi^{(n)} \right|^2 \to 0 \quad \text{as } n \to \infty.$$

Applying Proposition 6.1 to $(A^{(n)}, \phi^{(n)})$ on the space-time cylinder $\widetilde{K}^j_{[t,x]}$, we obtain a limit $(A_{[t,x]}, \phi_{[t,x]}) \in \mathcal{X}^w(K^j_{[t,x]})$ (up to gauge transformations and passing to a subsequence) on a smaller space-time cylinder $K^j_{[t,x]} := (t - \sigma^{-1}r(j), t + \sigma^{-1}r(j)) \times B_{\sigma^{-1}r(j)}(x)$, which is a weak solution to (MKG) obeying

$$\iota_{X_0} F_{[t,x]} = 0, \quad \left(\mathbf{D}_{[t,x]X_0} + \frac{1}{\rho} \right) \phi_{[t,x]} = 0.$$

The cylinders $\left\{ K^j_{[t,x]} \right\}$ for $j \in \{1, 2, \ldots\}$ and $(t, x) \in C^2_{[2,\infty)} \cap C_j$ form an open cover of the cone $C^2_{[2,\infty)}$. By Lemma 6.15, the weak solutions $(A_{[t,x]}, \phi_{[t,x]})$ form weak compatible pairs on $\left\{ K_{[t,x]} \right\}$. Then by Proposition 7.3, these pairs are equivalent to a set of smooth compatible pairs in some refined open cover of $C^2_{[2,\infty)}$, which in turn is equivalent to a single global smooth pair (A, ϕ) on $C^2_{[2,\infty)}$, thanks to the fact that $C^2_{[2,\infty)}$ is contractible. By construction, the pair (A, ϕ) satisfies the following properties:

- The pair (A, ϕ) is a smooth solution to (MKG) obeying the self-similarity condition

$$\iota_{X_0} F = 0, \quad \left(\mathbf{D}_{X_0} + \frac{1}{\rho}\right) \phi = \frac{1}{\rho} \mathbf{D}_{X_0}(\rho\phi) = 0.$$

- The following local convergences hold:

$$^{(T)} J_T \left[\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)}\right] \to ^{(T)} J_T[A, \phi] \quad \text{locally in } L^1_{t,x} \text{ on } C^2_{[2,\infty)}, \tag{8.63}$$

$$^{(X_0)} P_T \left[\widetilde{A}^{(n)}, \widetilde{\phi}^{(n)}\right] \to ^{(X_0)} P_T[A, \phi] \quad \text{locally in } L^1_{t,x} \text{ on } C^2_{[2,\infty)}. \tag{8.64}$$

We extend (A, ϕ) to a smooth self-similar solution to (MKG) on the whole cone $C_{(0,\infty)} = \{0 \le r < t\}$ by pulling back $(A, \rho\phi)$ along the flow of X_0. Note that (A, ϕ) has finite energy (again bounded by $\le E$), thanks to the local convergence (8.63). Hence by Proposition 7.2, it follows that $\mathcal{E}_{S_t}[A, \phi] = 0$ for every $t \in (0, \infty)$. However, this is a contradiction with (8.49) (in particular, for large enough t so that $S_t^{(1-\gamma_2)t} \subseteq C^2_{[2,\infty)}$) and (8.64). $\qquad\qquad\square$

References

1. Bahouri, H., Gérard, P.: High frequency approximation of solutions to critical nonlinear wave equations. Am. J. Math. **121**(1), 131–175 (1999)
2. Bourgain, J.: Global solutions of nonlinear Schrodinger equations. Am. Math. Soc. Colloquium **46**
3. Cuccagna, S.: On the local existence for the Maxwell-Klein-Gordon system in R^{3+1}. Commun. PDE **24**(5–6), 851–867 (1999)
4. Eardley, D.M., Moncrief, V.: The global existence of Yang-Mills-Higgs fields in 4-dimensional Minkowski space. I. Local existence and smoothness properties. Commun. Math. Phys. **83**(2), 171–191 (1982)
5. Eardley, D.M., Moncrief, V.: The global existence of Yang-Mills-Higgs fields in 4-dimensional Minkowski space. II. Completion of proof. Commun. Math. Phys. **83**(2), 193–212 (1982)
6. Grillakis, M.: Regularity and asymptotic behaviour of the wave equation with a critical nonlinearity. Ann. Math. **132**(3), 485–509 (1990)

7. Keel, M., Roy, T., Tao, T.: Global well-posedness of the Maxwell-Klein-Gordon equation below the energy norm Discrete Contin. Dyn. Syst. **30**(3), 573–621 (2011)
8. Kenig, C.E., Merle, F.: Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case. Invent. Math. **166**(3), 645–675 (2006)
9. Kenig, C.E., Merle, F.: Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation. Acta Math. **201**(2), 147–212 (2008)
10. Klainerman, S., Machedon, M.: Space-time estimates for null forms and the local existence theorem. Commun. Pure Appl. Math. **46**(9), 1221–1268 (1993)
11. Klainerman, S., Machedon, M.: On the Maxwell-Klein-Gordon equation with finite energy. Duke Math. J. **74**(1), 19–44 (1994)
12. Klainerman, S., Machedon, M.: Finite Energy Solutions of the Yang-Mills Equations in $\mathbb{R}^{3+1}$. Ann. Math. **142**(1), 39–119 (1995)
13. Klainerman, S., Machedon, M.: Smoothing estimates for null forms and applications. A celebration of John F. Nash, Jr. Duke Math. J. **81**(1), 99–133 (1996)
14. Klainerman, S., Machedon, M.: On the optimal local regularity for gauge field theories. Differ. Integral Equ. **10**(6), 1019–1030 (1997)
15. Klainerman, S., Tataru, D.: On the optimal local regularity for Yang-Mills equations in R^{4+1}. J. Am. Math. Soc. **12**(1), 93–116 (1999)
16. Krieger, J., Lührmann, J.: Concentration compactness for the critical Maxwell-Klein-Gordon equation (2015)
17. Krieger, J., Schlag, W.: Concentration compactness for critical wave maps European Mathematical Society. EMS Monogr. Math. (2012)
18. Krieger, J., Sterbenz, J.: Global regularity for the Yang-Mills equations on high dimensional Minkowski space. Memoirs AMS (2005)
19. Krieger, J., Sterbenz, J., Tataru, D.: Global well-posedness for the Maxwell-Klein Gordon equation in 4+1 dimensions. Small Energy (2012)
20. Lawrie, A., Oh, S.-J.: A refined threshold theorem for (1+2)-dimensional wave maps into surfaces. Commun. Math. Phys. (2015). doi:10.1007/s00220-015-2513-7
21. Lions, P.L.: The concentration-compactness principle in the calculus of variations. The locally compact case. I. Ann. Inst. H. Poincaré Anal. Non Linéaire **1**(2), 109–145 (1984)
22. Machedon, M., Sterbenz, J.: Almost optimal local well-posedness for the $(3 + 1)$-dimensional Maxwell-Klein-Gordon equations. J. Am. Math. Soc. **17**, 297–359 (2004)
23. Moncrief, V.: Global existence of Maxwell-Klein-Gordon fields in $(2 + 1)$-dimensional spacetime. J. Math. Phys. **21**(8), 2291–2296 (1980)
24. Oh, S.-J.: Gauge choice for the Yang-Mills equations using the Yang-Mills heat flow and local well-posedness in H^1. J. Hyperbolic Differ. Equ. **11**(1), 1–108 (2014)
25. Oh, S.-J.: Finite energy global well-posedness of the Yang-Mills equations on R^{1+3}: an approach using the Yang-Mills heat flow. Duke Math. J. **164**(9), 1669–1732 (2015)
26. Oh, S.-J., Tataru, D.: Local well-posedness of the $(4 + 1)$-dimensional Maxwell-Klein-Gordon equation at the energy regularity. Ann. PDE (2016). doi:10.1007/s40818-016-0006-4
27. Oh, S.-J., Tataru, D.: Energy dispersed solutions to the (4+1)-dimensional Maxwell-Klein-Gordon equation (2015). arXiv:1503.01561 (**under review**)
28. Rodnianski, I., Tao, T.: Global regularity for the Maxwell-Klein-Gordon equation with small critical Sobolev norm in high dimensions. Commun. Math. Phys. **251**(2), 377–426 (2004)
29. Selberg, S.: Almost optimal local well-posedness of the Maxwell-Klein-Gordon equations in 1 + 4 dimensions. Commun. PDE **27**(5–6), 1183–1227 (2002)
30. Selberg, S., Tesfahun, A.: Finite-energy global well-posedness of the Maxwell-Klein-Gordon system in Lorenz gauge. Commun. Partial Differ. Equ. **35**(6), 1029–1057 (2010)

31. Sterbenz, J.: Global regularity and scattering for general non-linear wave equations II. (4+1) dimensional Yang-Mills equations in the Lorentz gauge. Am. J. Math. **129**(3), 611–664 (2007)
32. Sterbenz, J., Tataru, D.: Energy dispersed large data wave maps in $2 + 1$ dimensions. Commun. Math. Phys. **298**(1), 139–230 (2010)
33. Sterbenz, J., Tataru, D.: Regularity of wave-maps in dimension $2 + 1$. Commun. Math. Phys. **298**(1), 231–264 (2010)
34. Tao, T.: Global regularity of wave maps. I. Small critical Sobolev norm in high dimension. Internat. Math. Res. Notices **6**, 299–328 (2001)
35. Tao, T.: Global regularity of wave maps II. Small energy in two dimensions. Commun. Math. Phys. **224**(2), 443–544 (2001)
36. Tao, T.: Global regularity of wave maps III. arXiv:0805.4666
37. Tao, T.: Global regularity of wave maps IV. arXiv:0806.3592
38. Tao, T.: Global regularity of wave maps V. arXiv:0808.0368
39. Tao, T.: Global regularity of wave maps VI. arXiv:0906.2833
40. Tao, T.: Global regularity of wave maps VII. arXiv:0908.0776
41. Tataru, D.: On global existence and scattering for the wave maps equation. Am. J. Math. **123**(1), 37–77 (2001)
42. Tataru, D.: Rough solutions for the wave maps equation. Am. J. Math. **127**(2), 293–377 (2005)
43. Uhlenbeck, K.K.: Connections with L^p bounds on curvature. Commun. Math. Phys. **83**(1), 31–42 (1982)